VOL. XIV OCTOBER, 1930 No. 5

The Journal of APPLIED PSYCHOLOGY

JAMES P. PORTER, EDITOR
Ohio University, Athens, Ohio
WITH A BOARD OF CO-OPERATING EDITORS

TABLE OF CONTENTS

M. EUSTACE BROOM. A Study of a Test of Ascendance-Submission.. 405
ALEXANDER C. CROCKETT. A Measure of Manual Ability..... 414
JAMES D. WEINLAND. An Objective Method for the Measurement of Attitudes..... 427
DOROTHEA MCCARTHY. A Study of the Seashore Measures of Musical Talent..... 437
VERNE R. ROSS. A Preliminary Investigation of the Effect of Radio Reception on School Achievement..... 456
T. H. CUTLER. The Effectiveness of Page Size in Magazine Advertising..... 465
KATE HEVNER. New Tests for Esthetic Appreciation in the Field of Music..... 470
WESTON A. BOUSFIELD. A Study of Motor Skill in the Free Hand Duplication of Geometrical Figures..... 478
W. E. MCCLURE. The Status of Psychological Testing in Large City Public School Systems..... 486
L. E. RACKLEY. The Blood Pressure and Galvanic Reflex as Indicators of Emotional States..... 497
NOTES AND NEWS..... 505
BOOK REVIEWS..... 507
NEW BOOKS AND PAMPHLETS..... 514

PUBLISHED BI-MONTHLY BY THE EDITOR
All communications should be addressed to the Editor,
Ohio University, Athens, Ohio

Printed by
WAVERLY PRESS, INC.
Baltimore, Md.

Entered as second class matter June 3, 1924, at the post office at Baltimore, Maryland, under the act of March 3, 1879.
Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917. Authorized December 21, 1923.

COPYRIGHT 1930, BY JAMES P. PORTER

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University; W. V. Bingham, Personnel Research Federation, New York; June E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; Joseph Jastrow, University of Wisconsin; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry C. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Boston, Mass.; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; J. E. Wallace Wallin, Department of Education, Baltimore, Md.; G. M. Whipple, Danvers, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

A STUDY OF A TEST OF ASCENDENCE-SUBMISSION

M. E. BROOM

State Teachers College, San Diego, California

The purpose of this report is to present data concerning the Allport A-S Reaction Study, a test which purports to measure an examinee's tendency to dominate or be dominated by other persons. The problems considered include (1) a study of the reliability of the measuring instrument, (2) a study of its validity, (3) a study of sex differences discovered by the use of the test, and (4) a study of the use of the test, teamed with the Thorndike Intelligence Examination, Series of 1925–1929, for predicting scholastic success in college. The data used in these several studies were collected in the San Diego State Teachers College during the academic year of 1928–1929.

The reliability of the test is quite low. The coefficient of reliability is only positive .718 (P.E., .040, 68 cases). The value of this constant is slightly lower than the reliability coefficients reported by the test author.[1] The time interval between the two applications of the test was about three weeks in the present instance. This interval was probably sufficiently long to permit the subjects to forget their first reaction, but not so long that the validity of the results was seriously affected.

The validity of the test is a matter for some concern in view of its low reliability. The first step was to determine a criterion for validity. The writer undertook to meet this difficulty by constructing a simple rating scale for ascendence-submission similar to one he had previously used successfully for a similar purpose while studying reactions to the Laird Personal Inven-

[1] Allport, G. W. "A Test of Ascendence-Submission." Journal of Abnormal and Social Psychology, 23: 118–136. July–September, 1928.

tory C-2.[2] In the present instance the scale included a brief explanation of the meaning of ascendence and of submission and the linear scale for rating, as follows:

Submissive				(Neutral zone)						Dominant
1	2	3	4	5	6	7	8	9	10	11

Each subject was asked to rate himself truly as to his own tendency to dominate or to be dominated by others. At the same time, each subject was asked to rate each other person in the group with whom he was well acquainted. From two to eleven ratings were secured for thirty-six members of the group. Three weeks later the procedure was repeated as a surprise move since the subjects had not been warned that they would be asked to give second reactions to the scale. It seemed possible to ascertain through this procedure whether the subjects were able to recognize in themselves with any consistency any trait which they identified as being to some degree submissive or dominant. It was possible to determine, furthermore, whether their fellows recognized in them the same trait or traits in anything like the same degree and whether they could recognize this characteristic or group of characteristics consistently. As the subjects were seniors and juniors in college they had had numerous opportunities to form judgments about each other, in fraternity and sorority meetings, in athletic contests, and in extra-curricular and social activities.

The several series of ratings are summarized in table 1.

The measures of the reliability and of the validity of the rating scale are given in table 2.

The measures given in table 2 are high enough to indicate low reliability and a sufficient degree of validity for the rating scale. The subjects were able to recognize with marked consistency the presence of certain characteristics in themselves which they described as ascendence or submission (or some degree between those extremes). They were also able to recognize the same characteristics in their fellows to a high degree. It seems then

[2] Broom, M. E. "A Critical Study of a Test of Extroversion-Introversion Traits." Journal of Juvenile Research, 13: 104–123. April, 1929.

that the ratings may be used as a criterion for studying the validity of the test without introducing great errors.

The coefficients of correlation for test scores paired with ratings follow in table 3.

There is low agreement only between ratings and test scores representing measures of ascendence and submission. The

TABLE 1

Summary of ratings

MEASURE	NUMBER OF CASES	RANGE		MEAN	$P.E._M$	S.D.	C. of V.
		From	To				
First self rating	51	2	10	7.18	0.15	1.57	21.87
Second self rating	44	3	10	7.34	0.13	1.30	17.71
First average rating given by others	36	4	10	7.03	0.18	1.56	22.19
Second average rating given by others	36	4	10	7.17	0.17	1.48	20.64

TABLE 2

Measures of relationship between the several series of ratings

	SECOND SELF RATING	FIRST AVERAGE RATING GIVEN BY OTHERS	SECOND AVERAGE RATING GIVEN BY OTHERS
First self rating	.761 ± .043 (44)	.526 ± .075 (42)	.544 ± .079 (36)
Second self rating		.622 ± .069 (36)	.605 ± .071 (36)
First average rating given by others			.741 ± .051 (36)

measures in table 3 agree quite markedly with those given by the test author (see footnote 1). It seems obvious that a critical study of the validity of the test is to be desired.

It seems logical that a part of the study of the validity of the test should be a comparison of scores yielded by this test with scores yielded by other tests of personality and temperament

traits. The tests used include (1) Brotemarkle's Comparison Test, (2) Pressey's X-O Test for Investigating the Emotions, (3) Laird's Personal Inventory C-2 and (4) the Moss et al. Social Intelligence Test. The measures are presented in table 4.

The low correlations between the Allport test scores and the scores yielded by the remaining instruments are not surprising since the different tests purport to measure different things.

TABLE 3

Measures of the validity of the test against ratings as a criterion

	FIRST SELF RATING	SECOND SELF RATING	FIRST AVERAGE RATING GIVEN BY OTHERS	SECOND AVERAGE RATING GIVEN BY OTHERS
First A-S score....	.460 ± .074 (51)	.424 ± .083 (44)	.410 ± .094 (36)	.446 ± .090 (36)
Second A-S score..	.483 ± .076 (46)	.442 ± .084 (42)	.423 ± .092 (36)	.481 ± .086 (36)

TABLE 4

Measures of the validity of the test against the criteria, scores yielded by other tests of personality or temperament traits

	NUMBER OF CASES	r (PEARSON)	P.E.$_r$
Brotemarkle	95	−.038	±.069
Pressey X-O t. aff	89	−.034	±.071
Pressey X-O i. d.	89	.013	±.071
Laird C-2	101	.134	±.066
Moss	26	−.122	±.130

Perhaps the most significant thing is that of the negative coefficients. These measures, however, are not significant since they are always smaller than their probable errors. The coefficient between the paired Laird and Allport scores, although very low, tends to substantiate Bender's[3] finding of low but positive

[3] Bender, I. E. "Ascendence-Submission in Relation to Factors in Personality." Journal of Abnormal and Social Psychology, 23: 137–43. July–September, 1928.

correlation between ascendence and extroversion as measured by the Heidbreder test.[4]

The study of the sex differences discovered by the Allport A-S Reaction Study now claims our attention. One hundred sixty-one reactions given by women and 59 reactions given by men are summarized in table 5. The percentile distribution of scores is given in table 6.

TABLE 5

Summary of distributions of scores for men and for women

	NUMBER OF CASES	RANGE		MEAN	$P.E._M$	S.D.	C. of V.
		From	To				
Men	59	−41	57	10.1	1.73	19.76	195.64
Women	161	−46	63	16.3	1.13	21.31	130.74

TABLE 6

Comparison of the percentile distribution of scores for 59 men and for 161 women

PERCENTILE	MEN	WOMEN
100	57.0	63.0
90	35.6	43.9
80	27.6	35.0
70	19.8	29.8
60	12.7	24.2
50	7.9	17.9
40	4.9	9.4
30	2.1	4.6
20	−5.6	−6.6
10	−13.0	−11.3
0	−41.0	−46.0

The coefficients of variation are of little significance due to the low value of the mean (which, in turn, results from the negative scores yielded by the test). The difference between

[4] Heidbreder, Edna. "Measuring Extroversion and Introversion." Journal of Abnormal and Social Psychology, 21 (1926): 120–134.

the mean scores for men and for women is 2.02 times the standard error of the difference. The chances are 98 in 100 that the true difference (the difference between the true measures) is greater than zero. The difference discovered is about 67 per cent of what it should be in order to insure a difference always greater than zero.

Summing up the data concerning sex differences we find that no evidence of definite sex differences has been given. The observed differences are of at least doubtful significance. A more detailed and inclusive study is necessary before final conclusions may be established.

The use of the test in predicting scholastic success is of some significance. It is becoming standard practice for guidance officers to include in a battery of measuring instruments some "tests which aim specifically at showing up the temperamental defects which interfere with mental ability and thus mar achievement."[5] In the present study the Allport test has been teamed with the Thorndike Ingelligence Examination, Series of 1925–1929. The reading section (Part III) of the Thorndike Intelligence Examination was also considered, since "in the high school and the university probably two-thirds of all studying consists essentially in reading."[6]

The several distributions of test scores and the distribution of average grade points, chosen as the index of academic success for this study, are summarized in table 7.

The normality of the experimental group in terms of a larger group (1233 cases) was investigated. The differences between the means for the two groups for the Thorndike total equivalent and reading scores are such as occur due to sampling. They are not significant. The difference between the means for the Thorndike total equivalent score distributions was only 1.19 times the standard error of the difference; the difference between the means for the Thorndike Intelligence Examination, reading

[5] Porteus, S. D. "Guide to the Porteus Maze Tests." Vineland, N. J., Extension Department of the Training School, 1924. Page 8.

[6] Starch, Daniel. "Educational Psychology, Revised Edition." New York. The MacMillan Company, 1927. Page 194.

section, only 1.72. In the former instance the chances are 88 in 100 that the difference is greater than zero, and the difference is about 40 per cent of what it must be in order to insure a difference greater than zero. In the latter instance, the chances are 96 in 100 that the true difference is greater than zero, and the

TABLE 7

Summary of the distributions of scores yielded by three tests and of average grade points per unit

	NUMBER OF CASES	RANGE		MEAN	$P.E._M$	S.D.	C. of V.
		From	To				
Thorndike total equivalent scores	202	24	102	62.7	0.73	15.47	24.67
Thorndike reading score	202	9	93	57.9	0.90	18.91	32.66
Average grade points per unit	195	−.28	2.77	1.77	0.03	0.54	30.51
Allport score	220	−46	63	11.2	0.99	21.72	193.9

TABLE 8

Measures of the relationship between the several paired series

	THORNDIKE READING SCORE	AVERAGE GRADE POINTS PER UNIT	ALLPORT SCORE
Thorndike total equivalent score	.794 ± .018 (202)	.397 ± .041 (194)	−.012 ± .047 (202)
Thorndike reading score		.442 ± .039 (194)	.043 ± .047 (202)
Average grade points per unit			.043 ± .047 (195)

difference is about 57 per cent of what it must be in order to insure a difference always greater than zero.

The first step in teaming the Thorndike and Allport measures was to ascertain whether the instruments measure the same variables in the same amounts. The coefficients of correlation for the several series of measures when paired are given in table 8.

It would seem that the next step would be applying the

technique of partial and multiple correlation to evaluate further the measuring effect of the several instruments, with the computation of regression equations. A simpler although rougher technique has been used instead. The procedure merely involves the distribution of average grade points per unit according to the location of the subject in one of the four quadrants established by placing the median Allport score on one axis and the median for the Thorndike Intelligence Examination series of scores along the other. The data for the four frequency distributions thus established are summarized in table 9.

TABLE 9

Results of teaming the Allport and Thorndike Intelligence Examination reactions in predicting academic success

Allport	Thorndike Intelligence Examination Median: Below	Thorndike Intelligence Examination Median: Above
Above	N = 49 M = 1.57 SD = 0.53	N = 54 M = 1.75 SD = 0.49
Md		
Below	N = 48 M = 1.27 SD = 0.53	N = 44 M = 1.22 SD = 0.45

It seems then that the Thorndike Intelligence Examination and the Allport A-S Reaction Study, measuring quite different things as they do, may be teamed to predict academic success for the dominant type of personality. It is easy, furthermore, to predict quite accurately within the dominant group on the basis of mental ability, roughly, at least, that a member of this group will excel submissive persons in academic endeavors. The personality trait of submission seems more important than mental ability in achieving success within the two lower quartiles

of the Allport test reactions. The submissive individual makes markedly lower academic records than the dominant individual on the average, and peculiarly enough mental ability does not seem especially important in making high marks among the submissive group.

In conclusion, the evidence indicates that the validity of the Allport test is probably satisfactory in view of the low reliability of the instrument. Improvement in reliability either by lengthening the test or otherwise, would probably tend to increase validity sufficiently. There are no appreciable sex differences discovered by the test. The test appears to be a valuable supplement to measures of mental ability in predicting academic success, particularly within the range of the upper half of the distribution of scores.

A MEASURE OF MANUAL ABILITY

ALEXANDER C. CROCKETT

Psychological Clinic, Detroit Public Schools

INTRODUCTION

The purpose of this paper is to describe a measure of manual ability. A task was designed to attempt to gauge some of the non-verbal manual elements underlying hand skill. Its use is to determine the general type of neuromuscular organization response of the individual nervous system. In so doing it serves as a convenient agent for (*a*), the observer in the psychological clinic, (*b*) for the industrial psychologist in the selection of skilled hand workers, and (*c*) for the teacher of manual arts courses in establishing an index of manual ability of students in training.

It may be given individually or to a group of 6 or 8 persons. It is made up of three different types of hand movements, (twisting, thrusting, and clasping), and yet is simple enough for a child of 5 years to do. Its range in scores is 326 points. The procedure has been standardized. The instructions are written out in detail, and time is allowed for practice. The task provides an opportunity for observing the spontaneous behavior of a subject. This report gives the first results from examinations of 1021 persons, 496 boys and 525 girls.

Nature of group tested. The persons tested averaged 17.4 years of age and 10.2 grades. The lowest age was 12; the highest 25. The lowest grade was 6 and highest 12. Most of these individuals were applying for trade training programs either in school or industry.

APPARATUS

The three parts of the test are housed in a large box with hinged cover. The first part entails the screwing of nuts onto

bolts and the placing of the bolts, heads down, in their respective holes. Nuts and bolts are contained in separate partitions of a shallow cigar box, $8\frac{1}{2}$ by $5\frac{3}{8}$ by $1\frac{1}{2}$ inches. There are 36 nuts and 36 bolts. The number of bolts of different sizes are as follows:

5 round head bolts	$\frac{1}{4}$ inch diameter and	$\frac{3}{4}$ inch	long
5 round head bolts	$\frac{5}{16}$ inch diameter and	$\frac{1}{2}$ inch	long
5 round head bolts	$\frac{5}{16}$ inch diameter and	1 inch	long
5 round head bolts	$\frac{5}{16}$ inch diameter and	$1\frac{1}{2}$ inches	long
5 round head bolts	$\frac{1}{4}$ inch diameter and	$1\frac{1}{2}$ inches	long
5 round head bolts	$\frac{1}{4}$ inch diameter and	1 inch	long
4 flat head bolts	$\frac{5}{16}$ inch diameter and	1 inch	long
2 carriage bolts	$\frac{1}{4}$ inch diameter and	1 inch	long

All nuts are square except the smallest sizes, which are hexagon shaped. The board containing the holes in which the bolts are inserted measures $8\frac{1}{2}$ by $5\frac{3}{8}$ by $\frac{3}{4}$ inches. The holes are distributed as follows: Two square-shaped holes are at the top. Then there is a row of 4 holes for the next largest bolts. Next are six rows with 5 holes in each row, the size diminishing gradually, with the smallest holes at the bottom. The depth of the holes is about $\frac{1}{2}$ inch. The board has a $\frac{3}{16}$-inch bevel around the four edges of the surface. The centers of the holes are spaced at 1-inch intervals for both the rows and the columns, with the exception of the 6 largest holes, which are symmetrically distributed in two rows across the upper quarter of the board. The box is placed at the right with the board next to it at the left. Box and board occupy the middle half of the large containing box.

Instructions. "Here is a large box containing 3 smaller boxes and a board with holes in it. There are 3 parts to the work, which you are to do. The first part deals with the two middle sections (the box with partitions and the board with holes); the second part deals with the boxes at either end of the big box, together with the blocks; the third part deals with the two strips of wood on the cover of the big box, together with the blocks. We will take up each part at a time, beginning with the first. Now turn to page 2."

PART I

"Are you right-handed or left-handed? If you are right-handed then always pick up a nut with your right hand. If you are left-handed then pick up a nut with your left hand. Try to remember this.

"You are to screw each nut on its proper bolt, just far enough so that it will not drop off. Then turn the bolt upside down and put it in the hole where it fits. Then do the next one.

"It is easier to start with the largest, and proceed to the next largest and so on to the smallest. However, you may do it any way you want, just so long as you use your right or your left hand in the correct way.

"The faster you work the better your score will be. You will have 3 trials for this part. Try hard to improve a little each trial. You will be told when to start and when to stop. Now read over this page again to make sure you understand. Then try a few nuts and bolts."

PART II

The second part consists of two deep cigar boxes and 80 blocks. The box that contains the blocks measures $8\frac{1}{4}$ by 5 by 2 inches (inside dimensions); the one that receives the blocks measures 8 by $4\frac{5}{8}$ by 2 inches. The small blocks are 19 by 19 by 19 mm., with an allowance of 1 mm. They are made of maple. Each weighs 4.5 grams. The boxes are placed at either end of the large box, the container being at the left, and the receiver at the right. The boxes are just 12 inches apart.

Instructions. "Do not pick up the end boxes for this part.

"You are to see how many blocks you can pack into the right-hand box, from the left-hand box, which now holds the blocks.

"Be sure to pack the blocks close against one another across the narrow end of the box in rows of 6 each, and without any spaces between. Stray blocks will not be counted in the score.

"Fill up the bottom of the box before starting a second layer. You will have three trials for this part. Try hard to improve a little each trial. You will be told when to start and when to

stop. Remember you may use any method you wish. Use both hands. Now read over this page again, to make sure you understand. Then try a few blocks."

PART III

The third part consists of two strips of cigar-box wood glued to the cover of the big box. Each strip measures 130 by 41 by 5 mm. There is a width allowance of 0.5 mm. The strips are glued $3\frac{5}{16}$ inches apart and $2\frac{1}{2}$ inches from the left-hand end of the box cover. The blocks used for part two are also used for part three.

Instructions. "Turn the big box around so that its cover is nearest to you.

"You are to pick up 3 blocks at a time and place them in a row across the strip of wood nearest the middle of the cover, and $\frac{1}{8}$ inch from the back of the cover. Then get 3 more and place them in the same way, $\frac{1}{8}$ inch away from the first row. Continue working on the strip until it is filled with rows of blocks and then start on the right-hand strip. Leave a space of about $\frac{1}{8}$ of an inch between rows, and at the top and bottom of each strip.

"Do not let the blocks fall off the strip. If one should fall off, replace it before going on with others. Be careful not to jog the table.

"When I say stop, take your hands off the blocks immediately. Remember to keep the rows about $\frac{1}{8}$ of an inch apart.

"You will have 4 trials for this part. Try hard to improve a little each trial. Now read over this page again to make sure you understand. Then try 4 or 5 rows of blocks."

The three boxes and one board that comprise the test are placed side by side loosely in a large box with a hinged cover. The cover has exactly the same dimensions as the large box. This box is 23 inches long, $9\frac{3}{4}$ inches wide and $1\frac{3}{4}$ inches deep. All box parts are painted a black lacquer. The small blocks and nuts and bolts are unpainted.

PROCEDURE

The subject is seated at a table whose top is 30 inches from the floor. The test is placed in front of him, and pulled toward him so that its front is flush with the edge of the table. He is given the manual of instructions and told to read it carefully. He is then permitted to practice the operations until he understands what he is to do, and has the "feel" of the material he is to use. He is given three trials for the first part, three for the second, and four for the third. The first part takes 2 minutes; the second, 1 minute; the third, 1 minute. The test requires 13 minutes actual working time. A rest interval of about 1 minute is allowed between each trial. The gross time required is approximately 30 minutes. The scores for each trial are recorded; those for each part are totalled and recorded. The grand total score is the sum of the three part scores.

RESULTS

Range in test scores. The lowest score made was 106; the highest 432. The range is 326 points.

Reliability. The last two trials of parts one and two, and the sum of the first two and sum of the last two of part three were correlated to determine the reliability of the test.

Part 1, trials 2 and 3, $+0.77 \pm .02$
Part 2, trials 2 and 3, $+0.77 \pm .02$
Part 3, trials 1 plus 2 and 3 plus 4, $+0.68 \pm .03$

The reliability could be increased by giving the subject less opportunity to choose his manner of work. If this were done, however, the diagnostic value of the test would be lessened.

Inter-correlations. Each part of the test was correlated with the other two parts with the following results.

Part 1 and 3 + .328 ± .06 N—125
Part 1 and 2 + .405 ± .06 N—125
Part 2 and 3 + .326 ± .06 N—125

There is a definite tendency for persons who score high on part 1, to score high on part 2, but they may score high on part 2 and low on part 1. Part 1 is, therefore, a more severe test

than part 2, and tests a more complex function. There is a definite tendency for persons who score high on part 3 to score high on part 2, but those scoring high on part 2 may score low on part 3. Part 3 is, therefore, a more severe test than part 2. Individuals who score high on part 2 are likely to score high on part 3, and those scoring high on part 3 are likely to score high on part 1.

Part 1 comprises the ability in part 2 (speed in grasping and packing) plus something in addition (visual acuity and touch and joint sensitivity). Part 3 comprises the ability in part 2, grasping, plus something in addition (touch and joint sensitivity).

Relation to other factors. The relation of the test scores to intelligence is $+.14 \pm .09$ for 87 persons. The relation to schooling is $+.14 \pm .04$ for 250 persons. The relation to age is $+.16 \pm .04$ for 250 persons. The relation to the Detroit Mechanical Aptitude test is $+.31 \pm .07$ for 83 persons.

Sex differences. The median score for each trial, and for the total score for each of the three parts of the test, and for the grand total score were computed for 496 boys and 525 girls. Absolute differences (D) are shown in the lowest row of table 1.

Effect of practice. The task was performed by one individual for 26 days over a period of two months' time. Table 2 shows the scores made on the first day of practice and the average scores made for the 26 days of practice. The greatest gain is made in part II with an average difference of 12 points per trial. Part I shows an average difference of 6 points per trial. Part III shows the smallest gain with 3.5 points per trial. The average for the three parts is a gain of 7 points.

Validity. (1) Correlation with success in a shop. The piecework production records for 26 toolmaker apprentices were correlated with each of the three parts of the test and with the total score. Most of the coefficients shown in table 3 have large probable errors, but five are sufficiently high to warrant attention, viz., part 2 and the drill press, part 1 and the external grinder, part 2 and the hand screw, part 2 and the total piece-work wage, and the total test and the hand screw.

TABLE 1*

Sex differences in median scores

N = 525 girls and 496 boys

	PART I			PART II			PART III				TOTAL			
	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 4	Part I	Part II	Part III	Grand total
Girls	15	17	19	44	49	51	23	26	27	27	50	142	101	293
Boys	16	19	21	40	45	47	23	25	27	27	56	133	104	287
D.	−1	−2	−2	+4	+4	+4	0	+1	0	0	−6	+9	−3	+6

* For information about percentile norms with letter ratings and a special observation record consult author.

TABLE 2

Differences between first days scores and average of 26 days scores, and the average deviations for each of the 10 trials over 26 days

	PART I			PART II			PART III				AVERAGE
	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 4	
First days score..	23	22	25	55	64	61	37	35	34	36	39
Av. of 26 days.....	28	29	31	69	74	77	37	39	38	38	46
A. D. of 26 days...	2.3	2.0	1.9	5.3	7.8	7.9	2.3	3.0	3.6	3.2	3.9

TABLE 3

*Relation between manual ability task and success in production work for 26 apprentices over 32 weeks' time**

	DRILL PRESS	MILLING MACHINE	EXTERNAL GRINDER	HAND SCREW MACHINE	TOTAL
Part 1	−.06+.15	.09+.15	*.73+.07	.29+.14	.02+.15
Part 2	*.53+.11	.29+.14	.27+.14	*.56+.11	*.50+.11
Part 3	−.05+.15	−.34+.14	−.04+.15	−.14+.15	−.35+.14
Total	.29+.14	.11+.15	.18+.15	*.53+.11	.19+.15

* Significant relationships. When corrected for attenuation these coefficients became as follows: Part 2 and drill press, +.59; Part 1 and external grinder, +.83; Part 2 and hand screw, +.66; Part 2 and total, +.54; total test and hand screw, +.61.

The coefficients of reliability for the production records for the four machines are (*a*) +0.60 for hand screw, (*b*) +0.35 for milling machine, (*c*) +0.63 for external grinder, and (*d*) +0.67 for drill press.

(2) Correlation with speed in manipulative processes such as assembling and dissembling. Ten workers were judged for speed by a shop supervisor and were given the manual test.

Test score	Rank order	Rank order	Shop rating
380	1	2.5	Fast
350	2	2.5	Fast
340	3	2.5	Fast
320	4	5.5	Average
300	5	2.5	Fast
295	6	8	Slow
288	7	8	Slow
282	8	8	Slow
277	9	5.5	Average
200	10	10	Slowest

The relationship between the two measures is +0.89 ±.08, when the method of rank differences is used.

(3) Correlation with success in producing cigar boxes. Seven *experienced* girl employees were given the manual test. The relation between the test and boxes produced is indicated below.

Rank order	Test score	Rank order	Production boxes
1	374	1	400
2	356	2.5	300
3	330	2.5	300
4	307	5	200
5	284	5	200
6	251	5	200
7	227	7	100

The relation here is +1.00, when the method of rank differences is used.

(4) Correlation with success in handling linen. Two fast workers and two slow workers were picked by the management

and were given the manual test. The two fast workers scored 293 and 332 with an average of 312. The two slow workers scored 214 and 214 with an average of 214. The difference in average scores was therefore 98 points, or about one-third of the entire range of scores of persons tested.

DISCUSSION

Meaning of terms. There is some confusion over the use of terms describing ability or abilities called mechanical. For instance, Thorndike (5) speaks of mechanical intelligence. By this he means "the ability to learn to understand and manage things and mechanisms such as a knife, gun, moving machine, automobile, boat, lathe, piece of land, river or storm." Thorndike makes no distinction between the person who is inventive, such as Edison, and the person who is skilled in a motor sense and is able to drive out home-runs, as Babe Ruth.

Stenquist (6) describes the mechanically inclined person as "one who feels a genuine thrill and joy in the contemplation or operation of a machine *because of its intrinsic mechanical nature.*" Stenquist here recognizes (*a*) insight and (*b*) manipulation, and one is left to assume that the two go together. Thorndike's view is thus accepted by Stenquist. Hull (2) recognizes motor abilities, but points out the relatively low correlation between tests of *motor* abilities. He states "we may say that as a rule they (motor activities) correlate very low with each other; but they occasionally show rather high correlations. These facts seem effectually to preclude belief in a two-level theory of aptitudes." Muscio (3) whose investigation is mentioned by Hull found no significant intercorrelations among *motor* tests for 5 different groups comprising 88 subjects. He concluded that "tests for vocational selection for motor tasks must be made up of exactly the same motor activities as those comprising the tasks themselves and that there is no general motor capacity."

A French investigator, Ehinger (1), seems to have discovered a plausible reason for the disagreement among authorities. She experimented with 243 subjects who were given 4 mental

tests with an interval of three to eleven months between the first and second trials. Her results yield an average inter-correlation of .36. She compares this with Muscio's lower intercorrelations, and states that "it is possible that there does not exist "*capacite motrice generale*," and that there does exist a general manual ability (*habilete manuelle generale*). She explains the difference between the two concepts by saying that some of Muscio's tests would not be regarded as tests of *manual ability*, such as aiming, reaction time, muscular strength, pursuit pendulum.

The apparent controversy over the existence of a general mechanical intelligence, as originally mentioned by Thorndike, can be partially cleared up if we come to an understanding of terms. Let us think of (*a*) mechanical intelligence as inventiveness, or mechanical insight (spatial thinking); (*b*) motor ability as ability for any single unit motor task and as a relatively simple function; (*c*) manual ability as "expertness and grace in manual acts" and as a complex or general function.

The above discussion should illuminate the evidence provided by the results of the present investigation. Our intercorrelations averaged $+.35 \pm .02$. This points to the existence of some common factor or factors.

What is the nature of the common factor? Our data show that the test correlates very low with age, grade, and intelligence. The test is not a test of inventiveness, of information, nor of experience in the narrow sense of the word, to any noticeable extent. The tasks are so simple in themselves that experience tends to be kept constant, in the same manner that the use of nonsense material in examining the higher mental processes attempts to keep experience constant. The definition that seems to fit best is "native expertness and grace in manual acts," and this we have called manual ability. Speed and accuracy play a part; the sense organs play a part; as does volition.

The very slight differences between the scores of boys and girls is further evidence pointing to an ability distinct from inventiveness, which has seemed to be more widely distributed among males.

The low positive relation of manual ability with schooling is in agreement with common knowledge of manual ability. Dull pupils may work skillfully with their hands; bright ones may be clumsy.

The low relation of the task with intelligence (as measured by the Detroit Alpha) corroborates the low relation with schooling. In the manual ability task the language factor, comprehension, and other higher factors as imagination and memory have been nearly eliminated. In effect, many of the factors that comprise the usual intelligence test are absent in the manual task.

We have seen that certain investigators have found evidence of a general "mechanical" factor. Others have found no evidence of a general "motor" factor. Others have found evidence of a general "manual" factor. All three points-of-view are conceivably correct. To harmonize them we must first agree on the definition of the terms, which definition has been suggested above; and secondly, we must consider the difference between simple and complex activity. Let us take up this difference.

Hull describes a possible law of levels of ability. "The highly complex intellectual activities correlate highly with each other. The less conplex correlate with each other to an intermediate degree, and the relatively simple motor activities correlate with each other only slightly."

It would seem that Hull overlooks a certain emphasis upon the *complex* motor activity. It is true he quotes McFarlane, who "believes that her inter-correlation coefficients are so much larger than those of Perrin and Muscio because her tests involve more complex behavior." The results of our study confirm this latter view.

The manual ability test seems to mark the requirements that characterize an independent variable. The reliability is reasonably high. The inter-correlations are rather low, but positive. The correlations with other independent variables (age, schooling, intelligence) are low. The correlations with criteria are reasonably high. The criteria represent radically different shop situations.

The nervous system is in fact a system, an organization.

Our hypothesis is that the manual ability task is to a certain extent a measure of this system or organization in action. It gives us some sort of a definite picture of an individual at work. The parts that compose the whole of the task are representative of common manipulative activity.

If manual ability might be thought of as an independent variable, certain implications would follow. The "lazy" individual may not be quite so lazy as he is lacking in manual ability. The same is true of the "slow" individual. The "industrious" person may not be as industrious as he is gifted in manual ability. Thus our terminology would take on more specific and exact meaning. Our methods of handling people, and our selection of persons to do types of work might be made more intelligent because we would better understand each individual.

SUMMARY

1. A task for manual ability was devised, standardized, tried out, treated statistically, and validated.

2. The task seems to be a measure of a general factor, complex in nature.

3. This factor seems to be in the nature of an independent variable, relatively unrelated to age, schooling, and intelligence.

4. Four experiments in different work situations indicate that the test is reasonably valid.

5. The terms mechanical intelligence, motor ability, and manual ability should be treated as separate concepts.

6. Manual ability is shown as having some of the characteristics of a distinct trait, having for its basis the organization and functional activity of the nervous system.

BIBLIOGRAPHY

(1) Ehinger, Gertrude: Recherches sur le Developpement de L'Habilete Manuelle. Travaux de la section de Technopsychologie de l'Institut J. J. Rousseau. Dir. L. Walther.

(2) Hull, Clark L.: Aptitude Testing. World Book Company, Yonkers-on-Hudson, N. Y., 1928.

(3) Muscio, M.: Motor Capacity with Special References to Vocational Guidance. Brit. Jour. of Psychology, xiii, 1922.
(4) Schlieper, Freidrich: Der Entwicklungsgang einer manuellen Ge schicklichkeitslei stung. Ein Beitrag zur Klarung des Begriffes, Geschicklichkeits.
(5) Thordike: Intelligence and Its Uses. Harpers Magazine, Vol. 140, pp. 227–235, 1920.
(6) Stenquist, J. L.: Manual of Directions for Mechanical Aptitude Tests. World Book Co., Yonkers-on-Hudson, N. Y.
(7) Sherrington, C. S.: The Integrative Action of the Nervous System.

AN OBJECTIVE METHOD FOR THE MEASUREMENT OF ATTITUDES

JAMES D. WEINLAND

School of Commerce, New York University

Since intelligence tests have shown themselves valuable in their field but limited in their scope to a restricted portion of the personality, attempts have been made to measure the emotions and character traits. The present discussion describes a further method of objectively measuring certain personality traits or attitudes. The foundation of the method is the maxim that the average response to a situation is the standard for a particular group and that variations from that standard are significant. This is precisely the same maxim that is made use of in constructing intelligence tests, the normal eight-year-old child being the one who has received the average or typical score for eight-year-old children. In measuring attitudes, however, two differences from the intelligence testing program will have to be noted.

First, since it complicates matters exceedingly to designate right attitudes and wrong attitudes while it may be perfectly possible to measure different attitudes, the norms will not be based on the number that answer right or wrong. The norm will be merely the average response to a given situation. But since entirely unguided responses would be so differentiated as to have no average or preponderant response, choice responses are called for to numerous situations. With two possibilities to choose from the preferred or preponderant selection will become the standard response to that situation. One choice will, of course, not evince an attitude, but fifty choices or one hundred may, particularly when the choices of different people are contrasted according to some principle, such as sex, group uniformity, etc. And when the situations to which choices are

asked have been experimentally selected for their diagnostic value definite results can be obtained.

And second, since attitude is no more the consequence of ability than it is of preference, habit, and other factors, the norms should probably be established for different groups such as college students, doctors, lawyers, salesmen, etc., rather than for different age levels or ability levels. Besides aiding in the measurement of personality such a test may then help in vocational diagnosis and placement.

We attempted here to get some measure of masculine and feminine attitudes and of the attitude of conformity versus nonconformity. These attitudes were selected because they seemed more tangible and easy to work with than many others, and so the logical place to start.

First of all it was necessary to select a situation the responses to which would reveal some of the attitudes of the individual. For this situation we selected a list of proverbs to be considered as statements of opinion and responded to by the judgment of true or false.

The reason for selecting proverbs may be given in a few words. First, they are each one a statement of attitude or opinion that has been representative enough to lodge in people's memory. Second, they are very brief enabling the recording of many responses which of course is a necessity in obtaining reliability. By calling for many responses, chance differences will be eliminated by canceling each other out and the permanent trends will show through. Third, by asking for an opinion of the correctness of proverbs the true intent of the test can be hidden and self consciousness reduced. Fourth, many proverbs are so little agreed upon that it is not easy to see the "expected response," and even an intelligent person cannot be sure of agreeing with the norms and getting a predicted score. He will find sincerity in his responses as safe as insincerity.

The test, then, consisted in the list of proverbs given below with the request that they be marked true or false and the statement that we were making a study of opinion.

STATEMENT OF OPINION

Directions: If you feel that a statement is true mark it +; if you believe it is false mark it −. Please mark each statement either + or −.

Male ✓ Female.....Occupation: Clerk

(−) 1 Clothes make the man.
(−) 2 Do one thing at a time.
(+) 3 None but the brave deserve the fair.
(+) 4 Defeat is tonic to a brave man.
(+) 5 He who is determined has his work half done.
(+) 6 Quick decision is the mother of good fortune.
(−) 7 A poor excuse is better than none.
(−) 8 Facts are stubborn things.
(−) 9 Familiarity breeds contempt.
(−) 10 'Tis the farmer's care that makes the field bear.
(+) 11 There is no wisdom like frankness.
(−) 12 A gentle spirit makes a gentleman.
(+) 13 A good deed is never lost.
(+) 14 All happiness is in the mind.
(+) 15 Happiness invites envy.
(+) 16 One mad action is not enough to prove a man mad.
(−) 17 Listeners hear no good of themselves.
(+) 18 Absence makes the heart grow fonder.
(+) 19 What makes life dreary is the want of motive.
(+) 20 Like likes like.
(+) 21 The best always goes first.
(+) 22 Blushing is virtue's color.
(−) 23 All men are liars.
(+) 24 Blood's thicker than water.
(−) 25 Children are what the mothers are.
(+) 26 Big head, little wit.
(+) 27 Men usually believe what they wish to be true.
(+) 28 Men respect and women love.
(+) 29 Love is blind.
(−) 30 Man loves but once.
(+) 31 There is no place like home.
(−) 32 Almost everything that is great is done by youth.
(+) 33 Little things affect little minds.
(−) 34 Every sweet has its sour.
(−) 35 A woman is naturally born to fear.
(−) 36 It is the men who cause women to dislike each other.
(+) 37 Men make wealth, women preserve it.
(−) 38 Women are sooner angry than men.

(−) 39 All's fair in love and war.
(+) 40 One love drives out another.
(−) 41 Might makes right.
(−) 42 Many women, many minds.
(−) 43 Everything comes to him who waits.
(+) 44 A proud man is always a foolish man.
(+) 45 It is not well to have all we want.
(−) 46 Wealth rarely brings happiness.
(+) 47 Sense comes with age.
(+) 48 All battle is misunderstanding.
(+) 49 Cheerfulness is a duty we owe to all others.
(−) 50 An unhappy man makes a good man.

In the first test there were fifty male and fifty female subjects. This is not enough to give any high degree of reliability but it is enough to illustrate the method. The grading was as follows. The total number of plus and minus responses for the men and for the women was first determined. Then a plus answer was called a masculine response if more men than women had marked the answer plus. It was called a feminine response if more women than men marked the answer plus. So a plus or minus reply was marked masculine or feminine because of a large or small margin in favor of that interpretation; for we estimated that if this margin were due to chance it would soon be canceled, if it were due to a masculine or feminine attitude the addition of small margins in the individual score would soon make this somewhat certain.

It is apparent that to some proverbs the masculine-feminine disagreement would be more marked than to other proverbs. We tried evaluating this difference and scoring accordingly but found that it did not materially change our ratings while it increased the work considerably. We gave up this refinement in the present case though it could well be used, particularly in selecting proverbs by which the masculine and feminine attitudes would be brought out the more sharply.

We have then the plus and minus answers marked masculine or feminine as shown in table 1. The proverbs, of course, refer to those given above.

In grading any individual it remains but to substitute mascu-

TABLE 1

Plus and minus responses to the proverbs and their scoring

PROVERB	RESPONSES				SCORED	
	Men		Women		Male	Female
	+	−	+	−		
1	7	43	3	47	+	−
2	42	8	45	5	−	+
3	18	32	13 *2	35	+	−
4	43	7	38	12	+	−
5	48	2	46	4	+	−
6	18	32	9	41	+	−
7	8	42	2	48	+	−
8	35	15	29 *1	20	+	−
9	30	20	28	22	+	−
10	16	34	11 *1	38	+	−
11	20	30	33	17	−	+
12	12	38	20	30	−	+
13	44	6	48	2	−	+
14	43	7	38	12	+	−
15	31	19	27	23	+	−
16	48	2	40	10	+	−
17	33 *1	16	23 *1	26	+	−
18	29	21	24	26	+	−
19	41	9	40 *1	9	+	−
20	35	15	19	31	+	−
21	16	34	18	32	−	+
22	18	32	16	34	+	−
23	20	30	12 *2	36	+	−
24	43	7	43	7	+ or −	+ or −
25	8	42	2	48	+	−
26	7	43	4	46	+	−
27	42	8	42	8	+ or −	+ or −
28	22	28	26	24	−	+
29	39	11	31	19	+	−
30	8	42	7	43	+	−
31	43	7	39	11	+	−
32	10	40	7	43	+	−
33	32	18	36	14	−	+
34	32	18	36	14	−	+
35	19	31	15	35	+	−
36	21	29	26	24	−	+
37	9	41	17	33	−	+

TABLE 1—*Continued*

PROVERB	RESPONSES				SCORED	
	Men		Women		Male	Female
	+	−	+	−		
38	19 *1	30	22 *1	27	−	+
39	25	25	8 *1	41	+	−
40	34	16	18	31	+	−
41	9	41	7	43	+	−
42	30	20	39	11	−	+
43	8	42	6	44	+	−
44	15	35	13	37	+	−
45	44	6	37	13	+	−
46	23	27	29	21	−	+
47	32	18	21	29	+	−
48	31	19	34	16	−	+
49	47	3	43 *2	5	+	−
50	1	49	1	49	+ or −	+ or −

* Undecided.

line or feminine for his plus or minus score. This will give a ratio statement of his masculine or feminine attitude, as in the case of the following eight subjects:

Subject	*Masculine responses*	*Feminine responses*
Male	26	21
	24	23
	31	16
	18	29
Female	31	16
	17	30
	15	32
	21	26

* On three of the fifty proverbs the men and women were in agreement.

We have then in this test an objective ratio score of masculine or feminine attitude. It is worth mentioning again that in taking the test the subject can seldom estimate in advance whether his answer will be characteristic of the masculine or

feminine mind. In consequence the subject gains little by trying to give the "expected answer." This eliminates the greatest weakness in many self ratings and personality tests. The subject marks the test without being able to tell the meaning of his answers, since this meaning is dependent upon the way a multitude of other people mark their tests. Sincerity of response will be reasonably certain since insincerity brings no apparent gains.

THE MEASUREMENT OF CONFORMITY VERSUS NON-CONFORMITY

The writer made the same type of study a few years ago upon 91 college men using a different set of proverbs and grading them somewhat differently. The proverbs follow with the total number of plus and minus responses given to the right of each statement.

Mark the following sentences + if they are true, − if they are wrong. Mark each one either plus or minus. If you think them partly true and false, mark them plus or minus depending on whether you think the true or false element is the stronger.

*+	−		
46	44	( )	A loving heart is the truest wisdom.
20	71	( )	The most of men are evil.
67	44	( )	Amiability is the redeeming quality of fools.
40	51	( )	A lovely girl is above all rank.
28	61	( )	Thank heaven the female heart is untenable by atheism.
48	41	( )	A multitude of books distracts the mind.
15	74	( )	Chance governs all.
21	67	( )	A woman's fitness comes by fits.
61	28	( )	Happiness is not the end of life, character is.
59	31	( )	To the pure all things are pure.
25	64	( )	He that dies pays all debts.
63	26	( )	Selfishness is the base of all evil.
22	69	( )	Opportunity knocks but once.
8	81	( )	Luck is the largest factor in success.
18	72	( )	Any means are justifiable to a worthy end.
41	48	( )	Nine times out of ten your first judgment is right.
66	33	( )	A little knowledge is more dangerous than ignorance.

* The figures at the left give the plus and minus responses of 91 college men. In a few cases the questions were not marked either plus or minus.

52 37 ( ) The man who flatters a woman either hopes to find her a fool or make her one.

82 6 ( ) No man is great, nor will he be great, unless he is earnest.

86 5 ( ) Disappointment subdues the small man, but to the man of courage it acts as a spur.

76 14 ( ) Most of the good or bad qualities that men attribute to others can be traced directly to themselves.

60 29 ( ) There is much danger in having too many friends and not enough enemies. A man, for his own good, must have enemies.

44 45 ( ) Convictions are more dangerous enemies of truth than lies.

59 31 ( ) Men always endeavor to cast suspicion on persons whom they cannot endure.

60 30 ( ) He that lacks time to mourn, lacks time to mend.
Eternity mourns that. 'Tis an ill cure
For life's worst ills, to have no time to feel them.

78 13 ( ) Those who compare the age in which their lot has fallen with a golden age which exists only in imagination, may talk of degeneracy and decay; but no man who is correctly informed as to the past, will be disposed to take a morose or desponding view of the present.

53 34 ( ) Man's love is of man's life a thing apart;
'Tis a woman's whole existence.

74 15 ( ) No, the heart that has truly loved, never forgets,
But as truly loves on to the close;
As the sunflower turns on her god when he sets
The same look which she turned when he rose.

65 26 ( ) Beyond this vale of tears
There is a life above,
Unmeasured by the flight of years;
And all that life is love.

In scoring the individual man he was given the numerical value attached to his plus or minus answer; a figure arrived at by taking the number of individuals who marked a proverb plus or minus. For instance, if he marked number one plus he would be accredited with 46, because 46 men had marked it plus. If he marked two minus he would be given 71 since 71 men had marked this statement minus. The third question marked minus would give him 44, etc. The higher the total

for all his answers the greater the number of men with whom he had been in agreement, the lower his total score the more individualistic or non-conformist, to that group, he might be said to be. The 91 men were then put in order of merit for conforming to the group attitude. Of the men measured the three highest scores were 1229, 1199, 1195. The three lowest 908, 903, 804. The first three men were in agreement with many of their group, the last three with few.

Another way of measuring conformity or non-conformity with group attitude was tried the following year with 78 men. The following words and phrases were handed to them on slips of paper with the statement that they each one represented a value or asset in the living out of life. The men were asked to put them in order of merit, with the word or phrase considered representative of the greatest value at the top, the least valuable at the bottom and the others ranged between. Nearly the maximum span of items for an order of merit problem was used and all the words represented some accepted value so that the subject could not too easily guess at the "expected rating." The words with their average rating follow:

Honor	5
Honesty	7
Courage	8
Health	9
Ambition	10
Love	10
Loyalty	10
Intelligence	10
Sincerity	11
Fairness	11
Religion	12
Self control	12
Friends	13
Energy	13
Service to others	13
Sanity	14
Sacrifice	15
Kindliness	16
Sportsmanship	17

Work	17
Family of children	18
Sense of humor	19
Business connections	21
Wealth	21
Recreation	23
Beauty	23
Hobby	24
Luck	24
Social position	24

The individual was scored by comparing his rating with the average and totaling his deviations. For instance, if the individual rated "Honor" as 1 his deviation would be 4; honesty 12, his deviation would be 5; courage 2, his deviation would be 6, etc. Thus the individual who showed a low total deviation score would conform to his group opinion; one who deviated with a high score would be the non-conformist from that particular group. Here it will be noted the method of scoring is just opposite from the proverb test where a high score indicated conformity, and a low score individualism. With the method here used the scores run from 87 to 183, the 87 man deviating least or conforming best to the group opinion. With both the proverbs and the order of merit list of words the non-conformists could then be individually studied if that appeared desirable.

CONCLUSIONS

In its present state the report merely outlines a method. There seem to be many advantages, however, in the method. It opens up a way of objectively measuring parts of the personality not adequately measured hitherto. One of its greatest advantages is that the expected and general answers cannot be foreseen by the subject. Much more work is needed to select the best proverbs and words for use; to find how many responses are needed to make the test reliable, and to find out group differences and the possible use of such tests in vocational guidance. The tests as given do show group conformity and non-conformity and in the case of the masculine-feminine attitudes, the direction of the deviation.

A STUDY OF THE SEASHORE MEASURES OF MUSICAL TALENT

DOROTHEA McCARTHY

University of Minnesota

INTRODUCTION

One of the most serious and definite attempts to standardize a test for the diagnosis and prognosis of vocational aptitude is Seashore's battery of tests of musical talent. On the basis of a detailed experimental analysis of musical capacity into its component psychological elements he constructed the above-mentioned tests of sensitivity to various elements in music. Since they are in the form of phonograph records and are readily available and convenient to administer these measures have been quite widely accepted in the field of vocational guidance and by administrators in music schools; often rather uncritically with the belief that they are reliable and valid measures of musical talent.

As in the case of all tests of vocational aptitude the research worker is confronted with many difficulties in establishing the reliability and the validity of his tests. No test should be accepted whole-heartedly as measuring what it is said to measure without abundant experimental evidence as to its reliability and its validity and our interpretations of the test results should be made only in the light of such findings.

PREVIOUS LITERATURE ON RELIABILITY AND VALIDITY

Heretofore, little experimental evidence has been available on these points and in fact, at the time this study was under-

taken,[1] no evidence had been published in the literature about the use of these tests. Seashore does not present any material on reliability, and the evidence on validity is rather unsatisfactory. Brennan (1) recently published some validity coefficients using as a criterion the ranking by expert judges on the basis of musical performance. These correlations were low, ranging from .17 to .47 for the six tests used, with four of them falling below .30. Brown (2), using rank by music instructor on "natural talent," reports even lower correlations, all but one being below .20. McGinnis (7), using a similar criterion, found correspondingly low validity coefficients for the tests when applied to preschool children. Stanton (20, 21) gives rather crude reports from the Eastman School of Music, indicating a tendency for those whose scores on the tests are high to remain longer in the school while those with low scores tend to drop out. These reports, however, are quite unsatisfactory from the point of view of the statistical validity of the tests.

Brown (2) also gives figures for reliability based on retest after a four-months interval which indicate very low reliabilities for the test which he used. Somewhat higher reliability coefficients are reported by Griffitts (5) for retest after an interval of one week. Heinlein (6) reports raw data for the consonance test which he says indicates a very low reliability for that test. Recently, Farnsworth (3) reported on the reliability of the pitch, intensity and time tests using different types of split-test techniques, and then testing out the use of the Spearman-Brown prophecy formula. The coefficients which he reports run

[1] This study was suggested by the late Dr. W. S. Foster of the department of psychology at the University of Minnesota, and was undertaken by students in the course in Advanced Experimental Psychology under his direction. Those responsible for the collection of the data, and for some of the statistical treatment are T. A. Langlie, F. Foote, Marie Ness, and Dorothea McCarthy. Due to the untimely death of Dr. Foster there has been considerable delay in the publication of the results. The data have been entirely reworked by the writer and thanks are due Dr. Florence L. Goodenough and Professor Donald G. Paterson for advice on the statistical treatment of the results and for criticism of the manuscript.

around .50 for the intensity and time tests and about .72 for the pitch test. Thus we see that our knowledge is in a vague state regarding the validity and the reliability of the Seashore tests and the relative importance of the various factors measured by them in making up general musical talent. The present investigation is an attempt to study the Seashore tests from these points of view and to try to gain additional information as to the reliability and the validity of the measures.

METHOD AND PROCEDURE

In this study there were 164 subjects: 93 university students, and 71 children in the fifth and sixth grades. The adults were members of two sections of the elementary laboratory course in psychology at the University of Minnesota, with 31 in one section, and 27 in the other, and 35 were members of the Vocational Psychology class. Children in the fifth and sixth grades of the Washington School were selected for study after consultation with the Supervisor of Public School Music in the Minneapolis Public School System, because of the presence of unusually complete and supposedly accurate ratings of musical performance in that particular school.[2]

The laboratory classes were given the pitch, intensity and memory tests four times. Periods of two hours each on two different days were required for the tests. On the first day each test was given twice, in the order: pitch, intensity, memory, memory, intensity, pitch. Two days later, each test was again given twice, in the order: memory, intensity, pitch, pitch, intensity, memory. The 35 students in the Vocational Psychology class were given one administration of each of the five tests, pitch, intensity, memory, consonance and time. The children were given the four tests, pitch, intensity, memory and consonance. In all 108 children were tested, but on account of absences on different days of examination, and on account of

[2] I take this opportunity to thank the director of music in the public schools of Minneapolis, Mr. T. P. Giddings, and the principal of the Washington School, Miss Gertrude Coleman, as well as the teachers for their hearty coöperation in this study.

failure to follow the directions for recording the responses, only 71 complete records were obtained. Directions were followed well on all but the memory test, on which there were 14 who in spite of our precautions gave records which were improperly filled out.

In administering the tests to the adults the instructions in Seashore's Manual (12) were carefully followed. The tests were given under ordinary classroom conditions and the subjects were cautioned to maintain quiet throughout the experiment. The instructions in the manual were followed in giving the tests to the children also, but the language was simplified as much as possible and there were demonstrations on the blackboard of how to record the judgments. The instructions were repeated at intervals during the tests and the positions on the record blank were checked at the beginning of each new column of ten judgments. Since the school in which the tests were given was in a business district many traffic noises had to be contended with, but the children were well habituated to the noises from the street and probably were not markedly disturbed by them.

RESULTS

In the treatment of the data the percentage of right judgments on each test was found for each subject and these percentages were converted into percentile ranks according to Seashore's norms presented in his manual. The median per cent of correct judgments as well as the median percentile ranks, and the coefficients of variability for each group on each test are presented in table 1.

Inspection of this table reveals that all groups were somewhat above the average of the group upon which Seashore standardized his tests. The only instance in which one of these groups falls below the 50th percentile is that of the children's score on the memory test. It was noted above that this is the test in which the children found it difficult to follow the directions as indicated by the large number of records which had to be discarded because they were improperly filled out. This may

mean that this particular test is more in the nature of an intelligence test, a memory-span test, or a directions test at this age level. The coefficients of variability are presented here to aid in the interpretation of all the correlations. They indicate that the groups are similar, not only in central tendency but in variability as well. Since there were negligible differences

TABLE 1

Data for the various groups

TEST	MEASURE	EXPT'L. CL., AM, N 31	EXPT'L. CL., PM, N 27	EXPT'L. CL., BOTH, N 58	VOC. PSY. CL., N 35	CHILDREN, N 71
Pitch	Med. Per Cent Right	83	81	82	81	57
	Med. Percentile Rank	63	56	63	50	51
	Coef. of Var.			.56	.55	.67
Intensity	Med. Per Cent Right	91	91	91	91	77
	Med. Percentile Rank	66	66	66	66	62
	Coef. of Var.			.50	.43	.51
Memory	Med. Per Cent Right	76	74	76	80	46
	Med. Percentile Rank	58	55	55	71	37
	Coef. of Var.			.48	.55	.59
Consonance	Med. Per Cent Right				74	64
	Med. Percentile Rank				78	62
	Coef. of Var.				.52	.51
Time	Med. Per Cent Right				75	
	Med. Percentile Rank				37	
	Coef. of Var.				.51	

between the two sections of the laboratory psychology course, both made up of upper-classmen in the university, and since the same tests were given to both groups under the same conditions, the two classes were combined in computing the correlations presented in the other tables of this article.

AGE DIFFERENCES

It will be seen from table 1, that as represented by percentile ranks the children are about average in all but the memory test and their scores do not appear to differ markedly from those of the adults. However, these scores obscure marked differences between the performance of the children and that of the adults. Comparison of the scores of the two groups expressed in per cent of correct judgments reveals that the children's scores are much lower than those of the adults and in the case of the pitch test the children do little better than chance. The percentile norms for children in the fifth and sixth grades are much easier than those for adults, and Seashore (12) also presents an intermediate series of norms for 8th grade children.

The findings of this investigation, and in fact of all other studies which have employed these tests including those of Seashore, hardly seem to support Seashore's contention that the capacities which these tests measure do not vary with age. His book (13) is full of statements to that effect; for instance he says, "The physiological limit of pitch discrimination does not vary with age." And again, "So far as the ear is concerned, the child at five can probably hear as well as he ever will, and one important element in hearing is pitch. If we test hearing without testing range of information or special skill, we find that the child can hear pitch as well as the adult. . . . . A child can tell as well as an adult which of two tones is the higher if he understands what is meant by higher." He admits finding age differences, however, in work with groups and of course the very fact that he finds it necessary to publish norms for children at different levels indicates that he finds great differences in performance. In another passage he says: "With a child who has a musical ear one may obtain a reliable measure as early as the age of five and even earlier. Mental dullness and possession of a relatively poor ear requires more maturity for the reliability of the measures. At least 90 per cent of all children can be measured satisfactorily by individual tests at the age of eight. . . . . The differences we find for age in the above records are accounted for by the handicap the younger child has

in responding to the conditions of the test in a group. When we test children individually and are in a position to simplify the conditions, we obtain practically the same norm for children as for adults." Still another statement is quoted by Seashore (13) from F. O. Smith, who says: " 'The amount in favor of the adult shown in all group statistics is amply accounted for by the difficulty in making a reliable test on the young and by their lack of information. This statement is based primarily on two lines of evidence,—the common occurrence of fine irreducible records among young children, and the character of the conditions which are ordinarily overcome by instruction and training.' "

This last contention about the unreliability of the tests when used with children is contradicted by the findings of the present investigation, for as will be shown in the section on reliability, the reliability coefficients obtained for the children are as high and even higher for some tests than they are for the adult groups. A second point is that in giving the test to the children individually, and in modifying the instructions and making them very much more detailed and specific, the conditions of the test are changed. The test is thus made easier and no direct comparisons of results obtained under different circumstances can be made. Seashore maintains that the difference is due in the main to the lack of understanding on the part of the children of the directions or of what is meant by the terms "higher," "lower," "weaker," "stronger," etc. Certainly children of fifth grade level, and especially of eighth grade level, know the meaning of these words and in nearly all cases the directions or filling out the blanks are understood by the fifth grade children. The memory test seems to be the only one with which they have any difficulty in following directions and this test does require a very long span of attention from these children. Children who fail to comprehend the directions either do not fill out the blank or mark it very incompletely and sometimes they become discouraged and cross out what they have marked. These cases, however, are quite rare. It seems safe to assume that the children who do fill out every blank entirely, understand the directions of the test and the meaning of the judgments they are

making. The results of this and of other investigations do not support Seashore's contentions and it seems that one of two things must occur; either these tests are measuring intelligence as well as musical ability, and that their relationship to intelligence or to mental age must be quite marked in order to account for the large age differences, or else that the capacities measured by the tests increase with age at least until maturity.

PRACTICE EFFECT

Since the tests of pitch, intensity and memory were each given four times to the 58 adults the data so obtained gave a good opportunity to test out some of Seashore's statements about the effect of practice on the results of various tests. Regarding the "Sense of pitch" Seashore says: "The sensitiveness of the ear to pitch differences cannot be improved appreciably by practice." As evidence for this he cites the work of Buffum (Psy. Mon. No. 53.) "who classified twenty-eight eighth-grade pupils in a preliminary fifteen-minute individual test for each one and then proceeded to train them by requiring forty minutes of the most specific and intensive practice under favorable conditions for twenty consecutive days, recording the average achievement for each day. The result was greatly surprising. Instead of showing progressive improvement with practice, each child remained as rated in the preliminary test, with only two exceptions, and the average ability for the class was the same at the end of twenty days of this intensive training as it was at the beginning." Seashore (13) also gives similar evidence in the case of the intensity test. He says: "Fourteen observers were selected regardless of initial ability and subjected to a prolonged series of specific and intensive practice of from fourteen to thirty periods of one half-hour each, and a record was kept of performance from day to day. The general conclusion from this experiment was that systematic practice under the most favorable conditions did not result in a change of the record. On the whole the achievement of these fourteen observers was practically the same at the end of the practice periods as it was at the beginning." He points to similar evidence

in the case of the blind, and in the case of musical training, and concludes that intensity discrimination and pitch discrimination are elemental capacities since they show no improvement with practice. Regarding the memory test, however, he says: "Memory is capable of cultivation on a large scale. The memory skill of any one in the plastic period might be increased at least tenfold, but the significance of records of this kind comes from the fact that they are marks of natural capacity. They show with what sort of raw material the memory training starts." He also states with regard to this test: "It may be said in general that the good records are probably reliable because they represent actual achievement. Ordinarily the upper half of the records should be considered entirely trustworthy."

Let us now consider these statements in the light of the findings of the present investigation based on four repetitions of these three tests. For this analysis the first and fourth administrations of each test have been compared as to central tendency and as to variability considering both the per cent right and the percentile ranks. Seashore advises using only percentile ranks based on the norms which he presents, but it was found necessary to use both types of analysis in order not to be led astray in the interpretation of the figures. Considering the median percentile ranks for these three tests it may be seen from table 2 that the results for the pitch test bear out Seashore's statement that the ability to discriminate pitch remains practically constant from first to fourth test. This also holds true when the median per cent of correct judgments is considered.

When the percentile ranks are used in judging the intensity test, however, there is an entirely different picture presented. There is an increase in central tendency from 66 to 82 in percentile rank from first to fourth administrations. This striking effect of practice in the intensity test is only apparent, however, for it is really a function of the measures used. Many people make very high raw scores on the intensity test and hence the norms expressed in percentiles decrease very rapidly so that a difference of one or two points in raw score makes a very large difference in percentile rank. When the per cent correct is used

as the measure and the medians and quartiles are found, the marked practice effect noted according to the first method disappears and there is a difference of only two per cent in the medians.

The results on the memory test show the improvement expected by Seashore, with an increase in median percentile rank from 55 to 75 and a corresponding increase in median per cent right from 76 to 84. According to Seashore (13), "such improvement in the use of a given brain capacity is no evidence of improvement in the capacity; it is nothing but evidence of acquired ability to use the available instrument or capacity."

TABLE 2

Medians and quartiles for first and fourth administrations

	PITCH				INTENSITY				MEMORY			
	Med.	Q_1	Q_3	Q	Med.	Q_1	Q_3	Q	Med.	Q_1	Q_3	Q
Based on percentile ranks												
1st 100	63	28	79	25.5	66	37	85	24	55	34	83	24.5
4th 100	63	32	87	27.5	82	51	87	18	75	41	91	25
Based on per cent correct judgments												
1st 100	82	75.5	85	4.75	91	85.5	93.5	4.00	76	66	88	11.00
4th 100	82.5	76.5	87	5.75	93	89	94	2.50	84	70	92	11.00

He further says, "The improvement will be roughly proportional to the original capacity; e.g., if we have the capacities represented by the 10 per cent, 50 per cent, and 90 per cent ranks respectively in three cases, the improvement is not going to make them any more alike than they now are, but will increase the difference in geometric ratio." This hardly seems consistent with his statement quoted above: "It may be said in general that the good records are probably reliable because they represent actual achievement. Ordinarily the upper half of the records should be considered entirely trustworthy." The figures on the variability for the memory test in the present

study indicate an improvement from the 34th to the 41st percentile rank for the lower quartile, and from the 83rd to the 91st percentile for the upper quartile with the semi-interquartile range remaining practically the same. This same tendency is also found when the per cent correct is used.

RELIABILITY

In this investigation the reliability of the measures of musical talent has been studied in two ways. The first method was by repetition of the tests on the experimental psychology classes. The various correlations between the different administrations of the three tests which were given four times to this group (table 3) are interesting from the standpoint of reliability, and

TABLE 3

Test-retest reliability coefficients

	PITCH			INTENSITY				MEMORY			
Expt'l. class N 58	2nd	3rd	4th		2nd	3rd	4th		2nd	3rd	4th
1st	.72	.73	.71	1st	.63	.65	.60	1st	.89	.85	.92
2nd		.75	.67	2nd		.59	.48	2nd		.87	.93
3rd			.69	3rd			.69	3rd			.94

for comparison with the odd-even reliabilities (see table 4), and also from the point of view of the stability of the tests. According to this method the memory test runs the highest in reliability even approximating Kelley's criterion for individual prediction. Pitch is next with reliability coefficients which are satisfactory for the use of the test with groups, and the intensity test has the lowest reliability.

One might expect to find in this table a tendency for increasing stabilization of the tests upon repeated administrations, indicated by lower correlations between the first and fourth administrations, and higher ones between the third and fourth administrations. The correlations however, are in such close agreement in most cases that there is no constant trend to be

noted. The differences which do appear among the intercorrelations for any one test are negligible and especially for the pitch and the memory tests they are well within the probable errors. The correlations between the first and second trials for these two tests are almost identical with those reported by Griffitts (5) and by Brown (2). The intensity test yields not only the lowest reliability coefficients by this method but also shows the most instability as indicated by the larger differences among the retest correlations.

The second approach to the problem of reliability was by the odd-even method. This method could also be used with the vocational psychology class and with the children. The per cent right of the odd-numbered judgments was correlated with the per cent right of the even-numbered judgments for each group. It was thought that this method of calculation might be rather misleading as it was in the case of the analysis for practice effect. In order to test this out several of the odd-even coefficients which it was thought would be most affected by the change were worked out using the per cent right of the odd-numbered judgments and of the even-numbered judgments. No constant tendency was noted for the correlations to be raised or lowered and the differences found were not large. Therefore, in table 3 the odd-even reliability coefficients are based on percentile ranks. For the group of 58 adults only the first administration of each test was treated in this way so that the figures are comparable with those for the other groups. All of these coefficients have been raised according to the Spearman-Brown formula

$$r = \frac{2r}{1 + r}$$

but both the raw and the corrected coefficients are presented.

These coefficients run quite high, three of them satisfying Kelley's criterion for individual prediction and four others approximating it very closely. The memory and pitch tests show high reliability according to this method also, and for all

groups. It must be remembered in interpreting these data that the memory and consonance tests consist of only 50 items whereas the others all require 100 judgments. In the split-test technique these tests are comparatively short and no doubt if they were as long as the other tests their reliabilities would be correspondingly higher. Of course, it would be possible with an extension of the Spearman-Brown formula to bring all of these reliability coefficients to the same standard and to calculate what the expected reliability would be were the tests all of equal length. It will be noted that when applied to children the tests seem to have a higher interest value for the children than for the adults. The latter usually complain of boredom while taking the tests whereas the children show great interest in the novel situation.

TABLE 4

Odd-even reliability coefficients

GROUP	N	PITCH		INTENSITY		MEMORY		CONSONANCE		TIME	
		Raw	Corr.	Raw	Corr.	Raw	Corr.	Raw	Corr.	Raw	Corr.
Exp. class	58	.77	.87	.62	.76	.74	.85				
Voc. class	35	.84	.91	.46	.63	.80	.89	.48	.65	.57	.73
Children	71	.86	.92	.82	.90	.75	.85	.42	.59		

VALIDITY

In order to secure a criterion with which to measure the validity of the tests the questionnaire (see table 5) was given to the adult subjects. It will be seen that this questionnaire is quite objective in that it asks for facts rather than for impressions or estimates. The experimenters re-rated the subjects on a scale of five as in question 8, table 5, using the factual answers to the preceding seven questions as a basis for their ratings. The mean of the self-ratings was 3.03, while that of the experimenters' ratings was 2.60, indicating a general tendency for these adults to underestimate their musical ability. There were 40 changes made by the experimenters in the ratings of the 93

subjects, only 8 of which were in the downward direction. Nevertheless, the correlation between the self-estimates and the experimenters' estimates was .87 when corrected by Sheppard's formula of correction for broad categories.[3] The experimenters'

TABLE 5

Questionnaire used with adults to obtain criterion of musical ability

1. Do you sing?	professionally .....
	before small audiences (friends, etc.) .....
	if others are singing also .....
	rather poorly .....
	not well enough to carry a tune .....
2. Do you play a musical instrument?	professionally .....
	only before small groups of friends .....
	only for own amusement .....
	just well enough to pick out a tune .....
	not well enough to pick out a tune .....
3. What instrument?	
4. Years of vocal training in music in school?	
Years of private vocal training?	
Approximate success, as indicated by grades?	
5. Years of instrumental training?	
6. Music courses in college with grades?	
7. Do you attend recitals, concerts, etc.	on every possible occasion? .....
	with fair regularity? .....
	occasionally only? .....
	practically never? .....
8. Give a general estimate of own general musical ability in one of the degrees indicated below:	
Rating of (1) if considerably above average	
Rating of (2) if somewhat above average	
Rating of (3) if average	
Rating of (4) if poor	
Rating of (5) if very poor	

ratings have been used throughout in the validity coefficients for the adult groups.

For the children, the music grades were used as the criterion with which to correlate the tests. All the grades were based on

[3] $\sigma = S - 1/12$. See Kelley, *Statistical Methods*, Sec. 47.

individual performance in singing before a special music teacher who had been with these classes for nearly a year. The letter grades were reported on a scale from A to F using plus and minus signs for the intermediate grades.

The percentile ranks on each test were correlated with these criteria in order to gain some idea of the validity of the tests. The results of these calculations, both raw coefficients and as corrected for broad categories by the use of Sheppard's formula are presented in table 6. It will be noted that these validity coefficients are considerably higher than those reported in the literature heretofore, especially those of Brown (2) and of Brennan (1). The memory test yields the highest validity coefficients among the adults. Pitch ranks next in validity,

TABLE 6

Validity coefficients

GROUP	N	PITCH		INTENSITY		MEMORY		CONSONANCE	
		Raw	Corr.	Raw	Corr.	Raw	Corr.	Raw	Corr.
Exp. class	58	.57	.63	.31	.34	.83	.93	—	—
Voc. class	35	.60	.66	.23	.25	.68	.75	.21	.23
Children	71	.43		.33		.40		.41	

and intensity gives the lowest correlation with the criterion. The very low correlation found for the vocational psychology class on the consonance test is in accordance with Heinlein's (6) work with this test in which he criticizes it very severely. The validity coefficients found with the children run lower than do those for the adults and they fail to bring out any differentiation among the various tests as was noted in the case of the adult groups. This, however, may be due to the difference in the criteria used with the different groups rather than to a true difference in the validity when used with children as compared with adults.

It is interesting to note that in this study as well as in those of Brown (2) and of Brennan (1), the memory test yields the high-

est correlations with the criteria used. Thus it seems safe to say that whether or not these criteria are adequate as measures of musical talent, they do give us something with which to compare the different tests one with another. Such findings must be due to some fundamental difference in the tests and it is quite striking that using such different criteria, three different investigations should indicate such similar results as to the relative standing of the different tests in validity as measured in this way.

TABLE 7

Intercorrelations

	GROUP	PITCH		INTENSITY		MEMORY	CONSONANCE
		Raw	Corr.	Raw	Corr.	Raw	Raw
Intensity	Exp. class	.24	.36				
	Voc. class	.23					
	Children	.25					
Memory	Exp. class	.56	.70	.43	.57		
	Voc. class	.41		.15			
	Children	.33		.20			
Consonance	Voc. class	.13		.14		.31	
	Children	.21		.06		.37	
Time	Voc. class	.25		.09		.04	.15

INTERCORRELATIONS

The intercorrelations between the various tests for the different groups are given in table 7. Those for the experimental psychology class for which retest reliability coefficients were available have been corrected for attenuation, and both the raw and the corrected coefficients are given in the table. In general these intercorrelations are low and bear out Seashore's (13) contention that the different tests measure different factors in musical ability. In several instances these correlations are almost identical with those reported by Brown (2) and most of them are in fairly close agreement with those reported by Seashore (13).

The highest correlations for all groups are those between pitch and memory. This is probably due to the fact that pitch discrimination enters into the memory test. When the subjects are asked to tell which of a series of tones is changed, it is the pitch of one of the tones in the series which is changed. The pitch discriminations involved in the memory test are not fine but this factor undoubtedly enters into the memory scores for those who are poor in pitch discrimination. Of the ten lowest persons in the experimental class in pitch discrimination eight are below the first quartile in the memory test.

CONCLUSIONS

In conclusion it may be said that the present investigation indicates first that there are marked age differences in performance on the Seashore tests which may mean either that the tests measure intelligence in addition to musical talent, or that musical capacity as measured by these tests improves with advance in age at least until maturity. There is no practice effect found after four repetitions of the pitch and of the intensity tests but there is a marked practice effect for the memory test.

As measured by the odd-even method the reliability of the pitch, intensity and memory tests is fairly high. Using the test-retest method the memory test shows the highest reliability, pitch is next, and the intensity test has the lowest reliability. The pitch and memory tests yield good validity coefficients when correlated with the criterion used for adults. In accordance with other investigations, this study shows low validity for the intensity test. The intercorrelations between the various tests run low and support Seashore's contention that the tests measure different factors. There is one exception to this, however, for the pitch and memory tests always correlate higher with each other than do any other two tests.

This study is not based on sufficient data to enable one to pass judgment on the consonance and on the time tests, but it appears that the intensity test is the least valuable for practical purposes. It has by far the lowest reliability and the lowet validity coefficients. The probable reason for this is the fact that the

test is too easy so that it does not differentiate well between members of a group, especially among those who are above the average. Probably the low scores on this test are the only ones which have any practical significance. The one test which is outstanding in both reliability and in validity is the test of tonal memory. This is especially surprising in view of the fact that this test is shorter than the others, having only fifty judgments. This test would probably prove the most useful single test in practical usage, not only because of high reliability and fairly high validity but also because of its relationship to the pitch test. In view of the facts pointed out above it might not be necessary to give both the pitch and the memory tests in many cases, for serious deficiency in the ability to discriminate pitch would be very likely to be revealed by the tonal memory test. This test, however, is rather difficult for children of fifth grade level. Probably the pitch test is the best single test at that age. In using these tests for individual guidance great caution should be used in interpreting the results. A good score in pitch is quite essential for one who expects to play the violin but this ability need not be possessed in such a high degree by the pianist. Likewise, the intensity test and the time and rhythm tests are probably of greater importance to the drummers and percussion members of an orchestra than to some other musicians. Repetition of the tests does not greatly improve their predictive value. The Seashore tests are of the greatest practical value at the extremes of the scale and as is the case with most measures of vocational aptitude it is much easier to predict failure than it is to predict success.

REFERENCES

(1) BRENNAN, FLORA M.: *The Relation Between Musical Capacity and Performance.* Psy. Mon., 1927, **36**, p. 190–248.

(2) BROWN, A. W.: *The Reliability and Validity of the Seashore Tests of Musical Talent.* J. of Appl'd. Psy., 1928, **12**, p. 468–476.

(3) FARNSWORTH, P. R.: *The Spearman-Brown Prophecy Formula and the Seashore Tests.* J. Educ. Psychol., 1928, **19**, p. 586–588.

(4) GAW, E. A.: *A Survey of Musical Talent in a Music School.* U. of Iowa Stud. in Psy., 1922, **8**, p. 128.

(5) GRIFFITTS, C. H.: *Fundamentals of Vocational Psychology*, Macmillan, 1924.

(6) HEINLEIN, C. P.: *An Experimental Study of the Seashore Consonance Test.* J. Exp. Psy., 1925, **8**, p. 408–433.

(7) MCGINNIS, ESTHER: *Seashore's Measures of Musical Ability.* Am. J. of Psy., 1928, **40**, p. 620–623.

(8) RUCH, G. M., AND STODDARD, G. D. *Tests and Measurements in High School Instruction.* World Book Co., 1927.

(9) SCHOEN, MAX: *Musical Talent and Its Measurement.* Mus. Quar., 1928, **14**, p. 255–282.

(10) SCHOEN, MAX: *The Validity of Tests of Musical Talent.* J. Comp. Psy., 1923, **3**, p. 101–121.

(11) SEASHORE, C. E.: *Elementary Tests in Psychology.* J. Ed. Psy., 1916, **7**, p. 81–86.

(12) SEASHORE, C. E.: *Manual of Instructions and Interpretations for Measures of Musical Talent.* Columbia Graphaphone Co., N. Y. C. Educ. Dept.

(13) SEASHORE, C. E.: *The Measurement of Musical Talent.* Silver, Burdett & Co. 1919.

(14) SEASHORE, C. E.: *The Measurement of Pitch Discrimination.* Psy. Mon. No. 53, 1911, **13**, p. 21–60.

(15) SEASHORE, C. E., AND LING, T. L.: *The Comparative Sensitiveness of Blind and Seeing.* U. of Iowa Stud. in Psy., 1918, viii, Psy. Rev. Mon., **25**, p. 148–158.

(16) SEASHORE, C. E., AND MOUNT, G. H.: *Correlation of Factors in Musical Talent and Training.* Psy. Rev. Mon. Sup., 108, 1918, **25**, p. 47–92.

(17) SEASHORE, C. E., AND TAN, K.: *The Elemental Character of Sensory Discrimination.* U. of Iowa Stud. in Psy., 1918, **8**, Psy. Rev. Mon., **25**, p. 159–163.

(18) SMITH, F. O.: *The Effect of Training on Pitch Discrimination.* U. of Iowa Stud. in Psy., 1914, vi, Psy. Rev. Mon., **16**, p. 67–103.

(19) STANTON, HAZEL, M.: *Inheritance of Specific Musical Capacity.* U. of Iowa Stud. in Psy., 1922, **8**, p. 57.

(20) STANTON, HAZEL M.: *Measuring Musical Talent: Seashore Tests as Administrative Aids.* The Personnel Journal, 1928, **7**, p. 286–292.

(21) STANTON, HAZEL M.: *Psychological Tests of Musical Talent.* Eastman School of Music, Univ. of Rochester.

A PRELIMINARY INVESTIGATION OF THE EFFECT OF RADIO RECEPTION ON SCHOOL ACHIEVEMENT[1]

VERNE R. ROSS

Covina High School California

The problem. What is the result on school achievement of the reception of radio broadcasts as received daily in millions of American homes? The perfection of radio apparatus and its wide distribution has led to a tendency for pupils during periods of home study to turn on the radio and carry on their study to its accompaniment. It is the purpose of this study to collect data bearing on this matter and to subject such data to statistical treatment in order to find an answer to the problem.

Scope of study. Radio effects on achievement in learned material and on the acquisition of new material were determined. In order that the experiment be controlled it was necessary to take it from the home and utilize recitation periods at school. The pupils used as subjects were two eleventh year classes in chemistry and one twelfth year class in physics, making a total of seventy-seven pupils. Arithmetic and reading were selected as the learned functions whose achievement was to be measured under the influence of the radio. Certain portions of physics and chemistry were used in that part of the experiment dealing with the acquisition of new material.

The achievement in learned material will be treated in Part I and in new material in Part II.

[1] Presented in May, 1929, to Educational Psychology Seminar of Dr. A. S. Raubenheimer, Department of Education, University of Southern California, Los Angeles, California.

PART I

Details of procedure. On coming to a certain class period the pupils were told that to succeed in physics or chemistry a thorough knowledge of arithmetic was necessary, and that it would be desirable to know at the beginning of the course just how much arithmetic they knew. They were then given a Stanford Achievement Arithmetic blank booklet and told to fill it out. The answers were checked in class and later rechecked by the instructor. At a succeeding recitation period the pupils were given Form B of the Stanford Arithmetic Achievement

TABLE 1

Mean scores and standard deviations of arithmetic and reading tests

	PHYSICS	CHEM. SEC. A	CHEM. SEC. B	TOTAL
IQ (Terman)	101	107	98	102
IQ (Sigma)	13	11	7	8
Arith. A	293	264	248	266
Arith. A (Sigma)	28	34	38	34
Arith. B (Radio)	285	260	247	263
Arith. B (Sigma)	37	28	34	33
Read. A	254	236		244
Read. A (Sigma)	26	23		24
Read. B (Radio)	251	239		244
Read. B (Sigma)	23	17		22

Test. During the time the pupils were working on Form B the radio loud speaker was in operation. During two other recitation periods Forms A and B of the Stanford Achievement Reading Test were given to the groups. Form B was completed to radio accompaniment.

Results of Arithmetic and Reading Tests. It is by a comparison of the mean scores of the groups as made on Forms A and B of the Arithmetic and Reading Tests that the effect of radio reception is to be determined. Table 1 gives the mean and standard deviations of Forms A and B of the arithmetic and reading tests.

Intelligence levels of the pupils. Included in the physics group

were two foreign pupils whose intelligence rating was lower than the ratings of the native pupils, due probably to language handicap. Also one other pupil had a low IQ. The average of the class exclusive of these three was 111. On the whole it was a decidedly superior group. The two chemistry sections were divided on an intelligence basis.

Discussion of data. In the arithmetic test the score on Form B made to radio accompaniment is less in each case than that on Form A made without radio. However, these differences are slight, ranging from one point in the case of chemistry, section B, to eight points for the physics group. In no case are these differences four times larger than their probable error. The reading tests show for the physics group a decrease of three points in Form B to radio accompaniment. Section A, chemistry, ranks three points higher in Form B to radio accompaniment. Section B, chemistry, was not given the reading test. These differences between the mean scores on the reading tests are less than four times their probable errors.

The variability of the scores as measured by the standard deviations are less on Form B with radio accompaniment than on Form A without radio accompaniment in all cases except that of the physics groups on Arithmetic Form B. In no case are the differences between the standard deviations of Forms A and B as large as four times the probable error of these differences.

Conclusion as to the effect of radio on arithmetic and reading. Since the difference between the mean scores of these groups of pupils on arithmetic and reading are not significant it seems that radio reception has no effect on performance in the learned functions of arithmetic and reading.

Items recalled from radio broadcast. While the radio broadcast was under way a log was kept showing the nature of the program being received. At the conclusion of the test the pupils were asked to list certain facts which they might be able to recall from the radio broadcast. These facts were demanded in the form of the following set of questions:

1. What time of day was announced?
2. What stations were on the air?
3. What was their location and telephone number?
4. What firms were advertising?
5. What were their addresses and telephone numbers?
6. What commodities were for sale?
7. What price were these commodities?
8. Name any persons mentioned.
9. State their connection with the broadcast.
10. List the names of any instrumental pieces of music heard.
11. List the names of vocal pieces of music heard.

A blank form containing the above set of questions was utilized by the pupils in listing what they could recall. A score was

TABLE 2

CORRELATION BETWEEN		PHYSICS	CHEM. SEC. A	CHEM. SEC. B	TOTAL
Arith. B	Radio Sc.	.24±.15	.38±.12	.17±.13	.32±.08
IQ	Radio Sc.	.52±.12	.37±.12	−.01±.14	.32±.07
Arith. A	Arith. B	.64±.02	.57±.02	.58±.03	.69±.04
Read. B	Radio Sc.	.47±.13	.42±.11		
IQ	Radio	.00±.17	.23±.13		
Arith. B (Radio)	IQ	.78±.06	.20±.13	.00±.14	.38±.07
Read. A	Read. B (Radio)	.88±.04	.77±.05		.82±.04
Read. B (Radio)	IQ	.79±.7	.75±.06		.74±.05

assigned to each pupil by allowing one point credit for each item listed which occurred in the broadcast. These scores were correlated with the intelligence quotients and with arithmetic scores made on Form B while the radio was operating. These correlation coefficients are shown in table 2. This table includes additional correlation coefficients.

The relation of radio scores and achievement scores. The correlation coefficients between arithmetic scores and radio scores are all positive low and not as large as four times their probable errors. Those for reading approximate four times their

probable errors. The values of r for IQ and radio score as made on the arithmetic test are significant for the physics group but not for the chemistry groups. In reading it is not significant for either group. The coefficients for the two reading tests are high and reliable. The coefficients for the correlation between Arithmetic B (radio accompaniment) and IQ are high and significant for the physics groups but of no significance for the two chemistry groups. Those for the physics group and chemistry section A are high and significant.

These results point toward some relation between intelligence and the ability to function successfully in reading and arithmetic in the presence of any distraction due to radio interference. The coefficients expressing the relation between reading ability and radio score indicate that those pupils who do well on the reading test also get the most from the radio program. The case of reading achievement and intelligence shows a high coefficient and indicates a close relation between success in reading and intelligence under the influence of the radio broadcast.

PART II

The experiment in detail. To determine the effect of radio reception on the acquisition of new material the three groups of pupils mentioned in Part I were utilized.

Before the experiment was begun all pupils were given a test on the material to be studied to determine what knowledge of the assignments was possessed. The text books of the classes were collected and kept in the laboratory the two weeks during which the experiment was conducted. Pupils were requested not to read outside material on the assignments. Ninety minutes per day on three days per week were available for the experiment. The pupils on coming to the laboratory were issued their text books together with an assignment. They were told to study this assignment for forty-five minutes, after which they would be given a test on the material studied and that the scores made on this test would enter into their quarter's grade. The physics group was divided into two sections, being paired so that the average intelligence of the two sections would

be equal. One of these physics sections together with one chemistry section studied for forty-five minutes with the radio operating. The other physics group and the remaining chemistry section were sent to a different room and studied for forty-five minutes without the radio. At the close of this study period both physics groups and both chemistry groups were given the same test on the material of the particular assignment.

The material of the assignments is shown in table 3. Preliminary tests on the material of these assignments were given. At the conclusion of the experiment the same tests were given again. It is by a comparison of the scores made on these preliminary

TABLE 3

Textual material of physics and chemistry assignments

ASSIGNMENT	PHYSICS		CHEMISTRY	
	Subject	Pages in Millikan and Gale Textbook	Subject	Pages in Newell's Chemistry
I	Change of State	173–84	Ammonia	148–57
II	Hygrometry	185–93	Nitric Acid	159–64
III	Boiling	193–203	Nitric Acid	165–70
IV	Steam Engines	204–10	Molecular Weights	171–76
V	Internal Combustion Engines	210–17	Molecular Weights	177–81
VI	Transfer of Heat	218–29		

and final tests that the effect of radio reception is to be determined. The two sections of the physics group were reversed as to radio study after the first three assignments. The two chemistry sections were treated in a similar manner. Both physics and chemistry groups were given that part of the original test which corresponded to the material of the first three assignments. On the completion of the last set of assignments they were given the remainder of the original test corresponding to the material of these assignments. The mean scores on these tests are given in table 4.

Findings on physics and chemistry material. Section A of

the physics group (with radio) in the study of the first three assignments gained 32 points while section B gained 38 points. This is a difference of eight points in favor of the non-radio group. On the second three assignments section A (without radio) gained 24 points while section B (with radio) gained 32 points. This is a gain of eight points in favor of the group studying to the radio accompaniment. In the chemistry group, section A with radio gained 62 points and section B without radio gained 58 points. The difference of 4 is in favor of the group studying with the radio. As the two groups were reversed as to radio study on the two last assignments, section A without

TABLE 4

Mean scores and gains made on physics and chemistry material.

SECTION	RADIO	PHYSICS GROUP			CHEMISTRY GROUP		
		Preliminary test	Retest after study	Gain	Preliminary test	Retest after study	Gain
A	Yes	34	66	32	21	83	62
B	No	29	67	38	24	82	58
A	No	27	51	24	5	20	15
B	Yes	19	51	32	6	18	12

radio is still making a greater gain of three points over section B who study with the radio.

These results clearly indicate that the reception of the radio broadcast is not a factor either facilitating or inhibiting the acquisition of facts of physics or chemistry.

Relations between intelligence and achievement in physics and chemistry. Coefficients of correlation were calculated between mental ages as determined by the Kuhlmann-Anderson Test, and the preliminary and retest scores for the various sections of the chemistry and physics groups. In no case were significant values obtained. Intelligence as measured in terms of mental age seems to play a small part in the success in physics study whether with or without radio accompaniment.

Results of individual assignments. At the conclusion of the study of each individual assignment each section of the physics and chemistry groups was given a test on the material of that particular assignment. A comparison of the scores made by the section studying with and without radio is shown in table 5. In the physics group section A leads by a small margin section B in five of the six tests, regardless of whether they are working

TABLE 5

Mean scores made on individual assignments by physics and chemistry sections

	RADIO	PHYSICS—MEAN SCORE	CHEMISTRY—MEAN SCORE
Assignment I:			
Section A	Yes	63	28
Section B	No	68	28
Assignment II:			
Section A	Yes	39	39
Section B	No	38	38
Assignment III:			
Section A	Yes	43	45
Section B	No	42	36
Assignment IV:			
Section A	No	43	36
Section B	Yes	40	36
Assignment V:			
Section A	No	44	38
Section B	Yes	42	27
Assignment VI:			
Section A	No	46	
Section B	Yes	43	

with or without the radio. This is the section which led in the gains on the retest over the preliminary. Other factors than radio reception are quite evident in the continual leading of this section. The differences in scores of the chemistry sections are small in three of the tests. In the test on assignment III, group A (with radio) shows a mean higher by nine points than group B, and on assignment V without radio the same group is eleven points ahead. These differences have a

probable error of 1.2 and 1.4 respectively, which indicate that the difference is significant. Since in one case this gain is made to radio study and in the other without it, no advantage or disadvantage can be assigned to the radio broadcast.

Items recalled from radio program. A log of the radio broadcast was kept and after a period of study and the test on the material of the assignment the pupils who had studied to radio were asked to list the facts which they could recall from the broadcast. Out of about 75 possible responses the maximum number recalled by individual pupils varied from zero to twenty-two. The average of the groups varied from three to nine. The greatest number of items recalled were made by individuals of high intelligence. Their scores on the assignments also ranked high. Coefficients of correlation between mental age and these radio scores were calculated. In no case were they large enough to be significant. Coefficients of correlation were calculated for the relation between mental age and the scores made on each individual assignment for the physics and chemistry groups. Some of these coefficients were fairly high but there was no regularity as to whether the high coefficients were with or without radio study.

FINDINGS AND IMPLICATIONS

1. Achievement in the learned processes of arithmetic and reading are not significantly affected by the reception of radio broadcasts.

2. The acquisition of new material in physics and chemistry is not significantly affected by the reception of radio broadcasts.

3. In that part of the experiment dealing with study of new material conditions under which it was conducted were such as to stimulate the pupils to do their best. It is doubtful if these conditions would be found in the home where individual study is carried on. Therefore these results should be interpreted as applying only to the conditions under which they were carried on.

THE EFFECTIVENESS OF PAGE SIZE IN MAGAZINE ADVERTISING[1]

T. H. CUTLER

Harvard University

With interest the public has observed the change in size recently of various popular publications. Some have enlarged, others reduced their former book or page size. To the scientifically minded this change has brought the question as to whether a page is a page no matter what the size may be. Does the maximum effort of the advertiser to get attention and induce the desired reactions succeed regardless of the dimensions of the page?

The problem is not to compare the relative value of two different sizes of advertisements in the same magazine, but to compare the large page with the small page in different publications. Will a full-page advertisement in a publication having a page size of four hundred fifty or four hundred twenty-nine lines have any less attentional or impressional value than the same full-page advertisement in a publication with a page size of seven hundred twenty-eight or six hundred eighty lines? If there is an advantage in size then the effectiveness of the page is dependent upon its size, and a page's value is according to its size.[2]

The value of a full-page advertisement compared with a half-

[1] I wish to express my thanks to Dr. R. H. Sylvester, Drake University, and Dr. E. G. Boring, Harvard University, for their cheerful help and criticism.

[2] The agate line, $\frac{1}{14}$ of an inch by 1 column, is the standard of measure for advertising space. The areas of the various standard size pages to which reference is made, are as follows: 728 lines, 117 square inches; 680 lines, 113.67 square inches; 450 lines, 73.4 square inches; 429 lines, 71.31 square inches.

page or quarter-page advertisement on the same size page has been quite satisfactorily proved.[3] However, the question arises as to whether the same factors that give the full-page advertisement the advantage over the half-page give an advertisement on a 728-line or 680-line page an advantage over the same advertisement in smaller copy on a 450-line or 429-line page. Any full page is admirable in that it develops prestige through its relation to advertisements of other sizes and possesses an isolation from other material. In these factors, however, the comparative full pages remain constant. So once more we face the problem: If an entire magazine is printed in a reduced size and everything in the magazine is proportionately reduced, would the attentional value of the advertisements be exactly the same in the smaller size as in the larger size?

II

The first task was to obtain the same advertisements in two different dimensions. Since the photostat process was found to be too expensive, magazines of two representative sizes—680 lines and 450 lines—were searched for duplicate advertisements. The advertisements selected were exact duplicates in copy material and illustrations. Proportional covers were used and dummies were constructed out of the thirteen advertisements that satisfied these requirements. The dummies were so constructed that the products were represented in the same order and position in both sizes.

The difficulty of measuring attentional value independently of memory value is well known. Recognition is more complete than pure recall. Consequently we have used both recall and recognition separately in the two parts of the experiment.

The experiment was presented to groups totaling three hundred ninety-three individuals. Both men and women, and high school and university students composed the groups. Each class of subjects was divided into two groups; one of these groups was given the small dummies, and the other the large.

[3] See any authority on psychology of advertising; e.g., D. Starch, Psychology in Advertising, 1925, Ch. VIII.

Two assistants aided the experimenter in giving the experiment. Preliminary instructions were first read and then the assistants passed out the material. The remaining details of procedure are evident from the instructions below:

INSTRUCTIONS

"This is a psychological experiment and hence we desire to keep as many factors constant as are possible. May I request that you do not turn the folders or pads of paper that will be given you without first receiving instructions and the signal to proceed.

"You now have folders representing the contents of a magazine. These folders contain a number of advertisements. You are to leave them on the desk before you and look over the advertisements on the right hand sheets as you might those which would interest you in a magazine. The speed at which you turn the pages, however, will be regulated. At the proper time interval I shall tap with my pencil and you are to turn the page promptly on that signal. To familiarize you with an interval I shall give you one now. (ten seconds) Please turn the pages carefully to avoid tearing and *observe only the right hand pages.* Are there any questions? Open—"

After inspection of the dummy:

"You are now to write on the first sheet of the pad of paper given you the trade names of as many of the advertisements as you remember. By *trade name* I mean 'the distinctive name given an article to distinguish it from all other commercial articles.' Do not guess! You will have four minutes in which to write these data. Are there any questions? You may proceed."

After the four minutes:

"On another sheet at the conclusion of these directions you will find a list of commonly advertised commodities. You are to mark a 'plus' at the left, which is just before the commodity which you recall as advertised in the folder you first inspected. It is desired that you first go down the entire list marking those you are positive of and then work at random. Are there any questions? There is but one sheet of the printed names, turn the pads over and proceed."

The method of scoring was based upon reproducing of the various trade-names. This method was chosen because it seems to approximate one of the most important advertising

situations. When the need or commodity is in mind and the advertised trade-name links itself therewith, then the desired association has been formed. Otherwise the advertisement has not been effective.

III

The results of the experiment consist of two scores obtained from each paper, one for recall and one for recognition. The average of 196 subjects for the small folder in recall was 6.75. The average for the same subjects in recognition was 10.65. Taking the averages of the values for recall and recognition as the "memory coefficient" we find 8.70 for the small size. With the large size and 197 subjects, we have an average score for recall of 6.97, for recognition of 10.63, and thus a "memory coefficient" of 8.80.

	RECALL	RECOGNITION	MEMORY COEFFICIENT
Small	6.75	10.65	8.70
Large	6.97	10.63	8.80

From these figures we find that the large size has the advantage of 0.22 in recall but that the small size has the advantage of 0.02 in recognition. In respect to the memory coefficient the large size has the advantage of 0.10.

Viewing these results as percents of the possible thirteen advertisements we find the same differences. In recall the large size has an avantage of 2 per cent while in recognition the percentages are equal. In the memory coefficient there is an advantage of only 1 per cent for the large size.

	RECALL	RECOGNITION	MEMORY COEFFICIENT
	per cent	*per cent*	*per cent*
Small	51	81	66
Large	53	81	67

IV

Such small differences lead us to the conclusion that increase in page size does not give corresponding increase in efficiency. When the small size is adequate for presentation of the copy, a larger presentation will not have any greater increase on attention or retention. The increase in value of the advertisement is not, therefore, proportionate to the increase in size and cost.

Tipper and Hollingworth believe in the operation of the law of diminishing returns, whereby the attentional value increases more slowly than the amount of space.[4] Poffenberger suggests that a square-root law may be active, namely that "changes in noticeable difference equals the square root of the changes in actual area."[5] Verification is also found in the article published while this work was in progress by Newhall and Heim, which concludes that retentional value is independent of absolute magnitude.[6] This is consistent with the contention of the *Gestalt* psychologists that perceptual size is not absolute. It may even be true that two pages of very different size may be phenomenally about the same when they are seen separately and not together. Whatever may be the difference there is agreement that absolute size has little practical importance.

Hence we must conclude that a page is a page provided that it is neither so small as to be difficult to read or so large as to be unruly. An advertisement in a publication having a page size of 680 lines will have little or no advantage over the same advertisement in a publication having a page size of 450 lines.

[4] H. Tipper and H. L. Hollingworth, Principles of Advertising, 1925, 100.

[5] A. T. Poffenberger, Op. Cit., 178–9.

[6] S. M. Newhall and M. H. Heim, Memory Value of Absolute Size, Jour. Appl. Psychol., 1929, 13, 62–75.

TESTS FOR ESTHETIC APPRECIATION IN THE FIELD OF MUSIC

KATE HEVNER

University of Minnesota

The great need for psychological tests in the field of esthetics is no doubt to be explained by the equally great difficulty of experimentation in that field. Esthetic qualities are not only elusive, but they are rather unstable, at least when present day measures of reliability are applied to them, and that "unerring good taste" which is so often claimed for certain individuals seems very difficult of demonstration.

The present experiment was an attempt to measure the appreciation of listeners for musical compositions played on the piano. It is different from other current tests for musical appreciation in that the materials for the test were actual musical compositions of various styles and types from the works of established composers and not bare patterns of rhythm nor lone sequences of chords. No attempt was made to analyze the beauty in good music and then to measure each of its component parts. Rhythm, consonance, melodic and harmonic progressions were not separated out and tested in isolation but they were left as the original composer had used them in his music combined and interrelated to build up the composition as a whole.

The test was composed of twenty-five items, each item constructed by selecting a composition by a writer of accepted worth and by changing or spoiling his work in three different ways. For example, one of the pieces was a series of eight measures from Chopin's Sonata in B minor, Opus 35. The subject listened to four versions of this music, played in succession. One version, called the "correct" was the composer's own transcript for the piano. A second version, the "eliminated" was

made by omitting many of the tones so that the music sounded bare, hollow and unexpressive, and a third version, the "elaborated," was made by introducing additional tones into the chords and by adding passing and auxiliary tones or ornaments. In a fourth version the phrasing or melodic progression was changed in one or two places. In every case the original work was followed as closely as possible so that the change should be in one direction only.

The items in the test were of different lengths ranging from three to sixteen measures although each one represented one complete musical phrase or idea. Examples were secured from three important musical periods, the classic, the romantic and the modern, a wide variety of styles and kinds of music was provided and each item was given an appropriate title suggestive of the kind of music the composer himself intended.

Nineteen different compositions were treated in this way and two other compositions were varied not by changing the compositions themselves but by changing the manner in which the pianist played them. Four of the items were repeated at some time during the presentation of the test which brought the total number of items up to twenty-five and the total time for the test to fifty minutes. The four different versions under each item were, of course, arranged so that each type appeared an equal number of times in the first, second, third and fourth places.

The items composing the test with the titles assigned to them and the number of measures used from each composition are as follows:

Composer	*Title*	*Title Assigned in Test*	*Measures*
Bach	Mass, B minor	Mass	9
†Beethoven	Symphony E♭ Major, Op. 55	Symphony	16
Bohn	Still as the Night	Song	4
†Brahms	Sonata F Minor	Sonata	4
Chopin	Prelude B Minor, Op. 28, No. 6	Prelude	4
Chopin	Sonata B♭ Minor Op. 35	Sonata	8

† Repeated later in the test.

Composer	*Title*	*Title Assigned in Test*	*Measures*
Chopin	Prelude A♭ Major	Prelude	3
DeBussy	Scherzo F♯ Minor	Scherzo	17
Godard	Au Matin, Op. 80	In the morning	3
Grieg	Ase's Tod, Op. 46, No. 2	Dirge	4
*Levitzky	Waltz	Waltzes	16
*Levitzky	Waltz	Waltz	16
MacDowell	By a Meadow Brook, Op. 51, No. 9	Meadow Brook	5
MacDowell	From Uncle Remus, Op. 51, No. 7	Humorous Piece	4
†MacDowell	In Autumn, Op. 51, No. 4	In Autumn	5
Mozart	The Violet	Song	3
Ornstien	Evening Prayer, Op. 9, No. 5	Prayer	5
†Ornstien	Song for the Hunters, Op. 9, No. 2	Hunting Song	3
Ornstien	March, Op. 9, No. 6	Ending for a March	3
Ornstien	Dance, Op. 9, No. 4	Dance	4
Schumann	Fantasie, C Major, Op. 17	Fantasie	13

* Varied only in interpretation of the pianist.
† Repeated later in the test.

The tests were played before four class groups at the University of Minnesota with approximately forty students in each group.[1] Two of the groups were composed of students in the Elementary Laboratory course in Psychology and two were groups of students in the Elementary Harmony course in the Department of Music. They were played by the same pianist on the same day, in the same place, and under conditions as closely identical as possible. The pianist had studied the music until her reading of it was perfect and her interpretation as constant as possible. The pianist was screened from the view of the listeners so that they could not get cues from her attitudes nor from the music script she was using. There was an interval of a minute and a half between the playing of each of the twenty-five items

[1] Grateful acknowledgment is made to Miss Margaret McKenzie and Miss Harriet Johnson, Seniors in the School of Music, who worked on this test as a part of their Laboratory course in Psychology especially to Miss Johnson who played the tests on the piano.

during which the subjects recorded their judgments and the experimenter announced the title of the next composition. The experimenter also added after the title, "This is a very short one," or "This one is rather long," or, "This is one of the longest items in the series," as the case might be.

Each listener received a mimeographed record sheet as follows:

TEST FOR INDIVIDUAL PREFERENCES IN MUSICAL COMPOSITIONS

Name..... Date.....

It is well known that individuals differ widely as to the kinds of music which they prefer. Quite different forms of musical composition have been developed by different races and in different periods of history. The various selections of piano music which you will now hear are designed to discover your own individual preferences in music.

You will hear some piano compositions in groups of four short pieces which are very much alike. Listen carefully and decide which of the four you most prefer, that is, which of the four sounds most pleasing, most appropriate, most "musical," to you. Write "Best" after the number which you found most pleasing, and "Second" after the number you liked next best. In some cases you may find it very easy and in some cases very difficult to make your judgments but in any case be sure that you mark one "best" and one "second" in every group. You may recognize the music in some of these groups and if so please write "familiar" in the margin following the numbers for that group.

The second page of the test blank contained a list of the twenty-five titles, and a space for recording judgments for each version. The papers were scored by allowing two credits for making the original or correct version "best," and one credit for marking it "second." The sum of all the credits earned constituted the subject's "score."

In comparing the scores of the seventy-eight harmony students with the eighty-one psychology students we would expect a slight superiority on the part of the harmony students and our experimental results do show this superiority. The mean score for the harmony students is 25.47 and the P.E. for the group is 3.57. The mean score for the psychology students is 23.7 and the P.E. 3.32. This gives a difference of 1.77 $\pm$.55 or 98 chances in 100 that there is a real difference between the groups.

A similar difference is noticed when all the subjects are grouped together and divided into quartiles on the basis of the test scores, for 17 per cent of the psychology students are included in the highest quartile and 31 per cent of the harmony students, and in the lowest quartile there are 30 per cent of the psychology students and only 20 per cent of the harmony students. That there is not a larger difference between the groups is probably due to a number of factors. (1) The psychology students were more accustomed to the testing situation and to the supervision of the experimenter and their attitude, attention, and effort were much better than those of the harmony students. (2) There were a number of well-trained musicians in the psychology group, as good as or better than many of the harmony students, and there were, likewise, some unmusical individuals in the harmony classes whose practical training in music was very limited and who were actually failing in the harmony course. The difference in score therefore is fairly satisfactory in the absence of a more complete differentiation between the two groups.

The 81 psychology students had previously classified themselves on a self-rating scale, for both musical talent, and musical training,[2] and their scores were also available on a variety of psychological tests. Correlations between the music appreciation scores and the scores on these several tests were calculated, and are as follows:

With Minnesota College Ability Percentiles	0.28
With Seashore Test for Consonance	0.204
With Seashore Test for Tonal Memory	0.236
With Seashore Test for Pitch	0.04
With Abbott and Trabue, Exercises in Judging Poetry	0.16
With Meier-Seashore Art Judgment	0.28
With Self-Rating of Music Talent	0.33
With Self-Rating of Music Training	0.30

It is interesting to note that overlapping of the new music test is about equally great in the fields of test intelligence, poetry and art appreciation and special aptitudes in music.

[2] See Foster and Tinker: Notes for Instructors to Accompany Experiments in Psychology, Revised Edition, p. 23.

An item-by-item analysis of the test brought out a number of interesting facts which will be of use in further developing and standardizing this type of test.

a. The repeated items did not show any consistency in the judgments of the listeners and invariably did not have as high diagnostic value as the original item. It is likely that they were often recognized as repetitions and that the new decision was attempted on the basis of memory. The subjects were certainly made confused and suspicious by their presence.

b. The items in which only the pianist's interpretation of the music was varied were not of any diagnostic value. These items were decidedly too easy for all groups. This type of variation is the most difficult for the pianist to present and it is possible that it might have a better diagnostic value if it were presented by means of an automatic player-piano so that more subtle changes could be made in the interpretation with the assurance that they could be kept constant throughout many presentations.

c. There does not appear to be any pronounced time error involved in presenting the four different versions in succession. The percentage of total judgments in favor of the compositions in the order in which they were played is as follows: First 22 per cent, second 26 per cent, third 24 per cent, fourth 27 per cent. There is a very slight advantage for the second and fourth places and a slight disadvantage for the version heard first.

d. The percentage of total judgments in favor of each of the four versions of the music was calculated and found to be as follows: For the correct 31 per cent, the elaborated 25 per cent, the changed phrasing 24 per cent and the eliminated 20 per cent. It is therefore obvious that the spoiled versions especially the elaborating and the changing of the melody or phrase were not sufficiently different from the correct version, nor sufficiently bad or unmusical. The most important revising of the test will be along these lines. It is also interesting to record that the psychology students showed a tendency to favor the eliminated or simpler version and the harmony students to

favor the elaborated version while the two groups showed an exactly equal tendency to prefer the changed phrasing.

e. In compiling the test an attempt was made to find music with established merit but with little familiarity for the average listener. Such items as Schubert's "Moment Musical," or Dvorak's "Largo from the New World Symphony" would be easily recognized and correctly identified by a great many listeners and were eliminated for this reason. The subjects were asked to mark as "familiar" any music which they recognized whether they could identify it or not and in the total number of judgments only 11 per cent were made on "familiar" material. Two-thirds of these were marked by the harmony students and one-third by the psychology students and more than half of them were made in connection with three items, Ase's Death, The Chopin Prelude in A Major and Bohn's song "Still as the Night." Of these 11 per cent marked familiar 0.057 per cent were votes for the correct version and 0.053 per cent for one of the spoiled versions. It is important therefore to eliminate familiar music as far as possible since not only does it constitute a source of error by introducing a different basis for judgment but it is a factor which tends to raise the scores of the listeners who are affected by it.

f. An analysis for the diagnostic value of the items was made in three different ways, by separating the talented from the untalented (self rating scales), separating the harmony students from the psychology students and separating the students with high scores from those with low scores. When the items were ranked according to their combined values in these three respects it is surprising that there is no clear cut difference in favor of the longer versus the shorter nor in favor of the classic versus the modern. Such a test as this may and should include music of widely various types and musical ideas of any length.

On the basis of these analyses a new scoring key was devised which eliminated eleven of the twenty-five items and which would reduce the presentation time for the test to one-half hour. All the repeated items were eliminated and the Levitzky waltzes which varied only in the performer's interpretation together

with the most familiar items and several others of low diagnostic value. Although the test was thus reduced to fourteen items the difference between the means for the harmony and the psychology groups remained equally significant and the correlations with the self-rating on talent was increased from .33 to .45, and with training from .30 to .42. The correlation between the odd and even numbered items which gives a measure of the reliability of the test was also raised very slightly.

On the basis of all the above results therefore, it has seemed worth while to work out a program for developing the test and plans are already under way for, (1) improving it by introducing more of the shorter and easier items in order to increase its reliability, (2) for presenting it to more sharply differentiated groups so that its exact diagnostic value may be known and (3) for examining the subjects who listen to it with other psychological and special aptitude tests so that its relation to other qualities and abilities may be discovered. The test may then be permanently recorded on an automatic piano and made available for popular use.

A STUDY OF MOTOR SKILL IN THE FREE HAND DUPLICATION OF GEOMETRICAL FIGURES

WESTON A. BOUSFIELD[1]

Harvard University

The original purpose of the investigation herein described was to study what appeared to be a simple, objective indication of motor ability and to correlate the results of a group test of this ability with intelligence and chronological age. Since the existence of motor ability as a functional unity has not yet been adequately demonstrated by these or other experiments, it is necessary to recognize the limitations imposed on an explanation of the nature of the work and on the interpretation of results. It seems advisable to avoid as far as possible the use of the term motor ability and to describe the process studied as one involving certain motor coördinations of a fairly specific nature.

Since Bryan's study of writing movement in 1892 and Bergstrom's experiments relating to card sorting in 1893, numerous attempts have been made to determine the relationship between various motor tests and to correlate the results with academic standing, intelligene and age.

Binet and Vaschide (1897) found a slight degree of correlation between simple and choice reaction time, rate of tapping, and speed of running but it was so low that they decided there was practically independence.

Bagley (1900) compared the scores made by children on one hundred sixty tests of motor ability and one hundred seventy-five tests of mental ability together with teachers' estimates and class standings. He concluded that there was a gradual

[1] The writer wishes to acknowledge an indebtedness to Dr. Donald Snedden of the Harvard Graduate School of Education for his kindly interest in the project, and numerous suggestions which have made the work possible.

increase of motor ability with age but there was at the same time a general inverse relation between motor and mental ability. Moreover, he found "numerous individual exceptions and varying validity at different periods of development."

In 1901 Wissler published the results of a correlation of certain mental and physical tests administered to Columbia Freshmen. Included among his tests were measurements of accuracy of movement, perception of time, association, perception of size, imagery and memory. These were supplemented with the available data concerning class standing. He decided: "(1) Laboratory mental tests show little intercorrelation in the case of college students. (2) Physical tests show a general tendency to correlate among themselves, but only to a very slight degree with mental tests."

The study made by Bolton (1903) would seem to give more favorable results. Motor ability he thought was divisible into a number of elements, the most important being: rapidity of voluntary control; steadiness and precision of movement; variety of actions; and quickness, strength and endurance of contraction. Various tests of these factors were given fifty-nine children from the lower wards of Lincoln, Nebraska, and fifty-two children from the best wards of the same city. Bolton says, "With brighter children, motor power increases with advancing age. There is a greater rapidity of motion, increased steadiness, and nicer precision the older the children grow. Backwardness, slowness of growth, and arrest of development are indicated by pupils through their inferiority to their fellows of the same age in some or all these respects."

Another division of motor ability is given by Ballard (1821). He says, "Motor ability, in fact, in any profitable sense of the term consists of at least three elements, strength, speed, and precision." The most careful researches, he believes have been made in England by Burt, Moore, Carey and Bickersteth. All of these investigators found a very small positive correlation between mental and motor ability, "an amount which got smaller and smaller as the subjects grew older and more intelligent."

Hanson (1922), on the basis of results obtained from an ingen-

ious device, concluded that the quality of continuous reactions after discrimination and choice, or serial action, furnished no reliable index of intelligence.

Glenn (1922) correlated a series of psychological tests with success in manual and academic activities. The average correlations between the different academic subjects varied from .60 to .80, whereas the correlation of the academic subjects with hand work gave coefficients varying from .12 to .42. These results were obtained from a study of children in the sixth, seventh and eighth grades.

The studies of Garfiel (1923) and Cowdery (1924) are in a sense related; for Cowdery adopted the main principles of Garfiel's technique.

Garfiel's tests were given to fifty college girls. She catalogued motor ability under five heads: (1) speed of voluntary movement; (2) accuracy of voluntary movement; (3) control of voluntary movement; (4) strength of voluntary movement; (5) motor adaptability. The intercorrelations of the motor tests varied from .15 to .25, while the correlation of the battery with the ratings of the ability of the subjects by three judges was .77. She writes, "Motor ability may be meaningfully understood, namely, that it represents a group of abilities in the main independent of, or at least different from mental ability." On the other hand, Cowdery, who attempts to refine Garfiel's tests, is led to the conclusion that "The interrelationships found raise definite doubt as to the existence of a distinct motor ability whose coördinated expression is relatively independent of general intelligence."

The recent work of Robert Seashore supports the evidence of low positive correlations between various motor tests.

A survey of the past investigations raises a doubt as to whether or not an adequate definition of motor ability has or can be formulated. The amount of functional unity between the various motor tests is still very much in question. For these reasons experimental evidence is theoretically valuable. The project under consideration has the aim of investigating a group of motor coördinations with reference to the dependence on intelligence and age.

A test composed of twelve parts was devised to secure an indication of the ability of children to duplicate with pen and ink various geometrical figures composed of simple curves and straight lines. The items were selected with the aim of combining a fairly comprehensive sample of simple geometrical constructions in a single test which could be administered to groups. The parts chosen were as follows: (1) three horizontal straight lines, $2\frac{3}{4}$ inches, 2 inches, and $1\frac{1}{4}$ inches respectively; (2) three vertical straight lines of the same lengths as the horizontal straight lines; (3) a right triangle with sides of $2\frac{3}{4}$ inches and $1\frac{3}{4}$ inches; (4) an equilateral triangle with sides of $2\frac{1}{4}$ inches; (5) three arcs of circles $2\frac{3}{4}$ inches, 2 inches and $1\frac{1}{4}$ inches in diameter, subtended by horizontal chords of 2 inches. The arcs only were given for duplication; (6) three arcs of circles of the same dimensions as (5), but placed vertically; (7) a circle 2 inches in diameter; (8) a spiral composed of two half circles, 2 inches and 1 inch in diameter; (9) a 2-inch square; (10) a cross composed of two 2-inch lines in the form of a large plus sign; (11) a reverse curve composed of two half circles 2 inches in diameter placed horizontally; (12) same as (11), but placed vertically.

The twelve parts were arranged on pages $8\frac{1}{2}$ by 11 inches so that the patterns were on the left half of the sheet. On the right half were points marked "O" from which all drawn figures were to begin. Three points were given for parts (1), (2), (5) and (6), one point for parts (3), (4), (7), (9), and (10), and two points for parts (8), (11), and (12). Where two points were given, one was designated by an "O" for beginning, and the other by a cross for ending.

The test was administered to one hundred eighty children of the fourth, fifth and sixth grades in two of the public schools of Brookline, Mass.[2] The children were urged to work carefully and to be as accurate as possible. Fifteen seconds were allowed for each part, signals being given regularly for starting and stopping. Most of the children completed the parts in about ten

[2] The giving of the tests was accomplished through the scientific interest and friendly cooperation of Principal Arthur W. Pearce of the Lawrence School and Principal Carlon W. Ray of the Heath School.

seconds. In order that greater consideration should be given to accuracy, it was announced at the beginning of each test that time would be allowed at the end for completion of the work if necessary. Only rarely was use made of the extra time allowance.

In order that the scoring of the test should be as objective and accurate as possible the errors were reduced to deviations in length, angle and area. This made possible the use of instruments, a cartometer, a protractor and a planimeter. The various parts were measured as follows: (1) Angular deviations were recorded by the protractor, and length deviations by the cartometer. For all representations of straight lines a summation was made of two measures: first, the distance of the end of the drawn line from the end of a true line at the same angle as the drawn line; second, the difference between the total length of the drawn line and a superimposed straight line. In this way a measure was obtained of the error of actual length estimation, and the irregularities of the line as it was drawn. (2), (3), and (4) were reduced to individual straight lines and treated in the same manner as (1). To estimate errors in (5) radii were drawn from the extremities of the drawn arc to the centre of the true circle. The true arc was superimposed, and continued if necessary until it cut the radii from the ends of the drawn arc. By means of the planimeter the area between the two arcs was measured. The difference between the angle subtended by the true arc and the angle subtended by the drawn arc was obtained by means of the protractor. (6) was treated in the same manner as (5). In (7) the area difference between the drawn circle and the true circle was measured by the planimeter. (8) was also measured by the planimeter. (9) and (10) were submitted to the same treatment as (1). (11) and (12) furnished area differences as (7). Errors in lengths were estimated to a 32nd of an inch, areas to hundredths of a square inch, and angles to a half a degree. In all, twenty sets of errors were obtained. Where the figures were scored on area alone a method of weights was used to make them equivalent to the rest. Means and standard deviations were calculated for each

TABLE 1

PART	LENGTHS		ANGLES		AREAS	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	0.91	0.58	7.86	5.91		
2	0.65	0.40	10.66	7.02		
3	0.98	0.54	10.54	5.39		
4	0.80	0.45	12.63	6.53		
5			16.70	10.02	0.95	0.38
6			22.06	14.40	1.12	0.44
7					0.83	0.49
8					0.80	0.47
9	0.61	0.27	9.63	4.88		
10	0.48	0.27	5.19	3.21		
11					1.04	0.54
12					0.72	0.36

Note: Lengths are given in inches, angles in degrees, and areas in square inches.

TABLE 2

	MEAN	S.D.
Scores on tests	49.50	3.57
Intelligence scores	32.59	4.62
Chronological ages in months	124.37	3.27

TABLE 3

Scores on tests of boys	49.71
Scores on tests of girls	49.29
Scores on tests of fourth grade	50.21
Scores on tests of fifth grade	49.24

Note: As all scores on tests were based on a summation of errors the higher scores indicate inferior ability.

TABLE 4

Correlation of scores on Drawing Tests with intelligence scores	$.180 \pm .069$
Correlation of scores on Drawing Tests with chronological age in months	$.015 \pm .071$

of the twenty series of scores. The results were then made comparable by equating the mean of each set of scores to 50 and adding one for each tenth of a standard deviation from the mean. For a final grade an average was taken of all the separate parts. In this way an equivalent weighting was given to the various items of the test.

Because of the large amount of time required to measure the various items it was thought advisable to omit the sixth grade. Of the remaining papers 90 were selected for complete scoring.

Intelligence grades were obtained from scores made by the children on the intelligence division of the Otis Classification Test which had been administered by the schools. Chronological age was figured in months from data supplied by the children on the test sheets. A summary of the calculations is given in tables 1 to 4.

CONCLUSION

The results of the study indicate a low positive correlation between the ability to copy accurately various geometrical figures and mental ability as indicated by the above mentioned intelligence test. The correlation of chronological age with the motor processes considered is entirely negligible in this range.

The differences in average scores on the drawing test indicate a slight superiority of the girls over the boys which is less, however, than the superiority of the fifth grade over the fourth grade.

A significant difference may be seen in the distribution of errors in the duplication of horizontal and vertical lines. Estimation of lengths is superior for vertical lines while estimation of angles is superior for the horizontal lines. This is clearly shown in parts 1 and 2 which consisted of separate horizontal and vertical straight lines.

Several criticisms may be advanced against the methods of the study. The selection of items for the test and the method of weighting is, to a certain extent, arbitrary.

In spite of attempts to secure a uniform timing for each individual it was evident that some children worked more rapidly than others and presumably sacrificed in accuracy. It is

possible that these discrepancies can be materially reduced by a general refinement of the test, the technique of administration and the method of scoring.

BIBLIOGRAPHY

(1) Bagley, W. C.: The Correlation of Mental and Motor Ability in School Children, Amer. Jour. Psychol., Vol. 12, pp. 193–205, 1900–1901.

(2) Ballard, P. B.: Mental Tests, pp. 196–209, Hodder and Stoughton, 1921.

(3) Bergstrom, J. A.: Experiments upon Physiological Memory by Means of the Interference of Associations, Amer. Jour. Psychol., Vol. 5, p. 356, 1893.

(4) Bolton, T. L.: The Relation of Motor Power to Intelligence, Amer. Jour. Psychol., Vol. 14, pp. 615–631, 1903.

(5) Brace, D. K.: Measuring Motor Ability, pp. 88–89. A. S. Barnes and Co., 1927.

(6) Bryan, W. L.: On the Development of Voluntary Motor Ability, Amer. Jour. Psychol., Vol. 5, pp. 123–204, 1892.

(7) Cowdery, K. M.: A Note on the Measurement of Motor Ability. Jour. Ed. Psychol., Vol. 15, pp. 513–519, 1024.

(8) Dearborn, W. F.: Intelligence Tests, p. 64, Houghton Mifflin, 1928.

(9) Garfiel, E.: Measurement of Motor Ability, Arch. of Psychol., No. 62, April, 1923.

(10) Glenn, I.: A Report on the Correlation of Psychological Tests with Academic and Manual Subjects, Jour. of Ed. Psychol., Vol. 13, p. 496, 1922.

(11) Goodenough, F. L.: Measurement of Intelligence by Drawings, World Book Co., 1928.

(12) Hansen, C. F.: Serial Action as a Basic Measure of Motor Capacity, Psychol., Monog., No. 140, pp. 320–382, 1922.

(13) Hull, C. L.: Aptitude Testing, pp. 300 ff., 359 ff., 413 ff., 477 ff., World Book Co., 1928.

(14) Jastrow, J.: A Sorting Apparatus for the Study of Reaction Times, Psychol. Rev., Vol. 5, pp. 279–285, 1898.

(15) Johnson, B. J.: Experimental Study of Motor Abilities of Children, Johns Hopkins University Studies in Education, No. 2, 1917.

(16) Koerth, W.: A Pursuit Apparatus: Eye-Hand Coördination, Psychol. Monog., No. 140, pp. 288–292, 1922.

(17) Perrin, F. A. C.: Experimental Study of Motor Ability, Jour. Exp. Psychol., Vol. 4, pp. 24–56, 1921.

(18) Wissler, C.: The Correlation of Mental and Physical Tests, Psychol. Monog., No. 16, pp. 1–62, 1901.

THE STATUS OF PSYCHOLOGICAL TESTING IN LARGE CITY PUBLIC SCHOOL SYSTEMS

W. E. McCLURE

Director, The Juvenile Adjustment Agency, Toledo, Ohio

Each new movement must demonstrate its right to a permanent place in our social order by proving that it is capable of filling an existing need. The need for psychological testing has long been recognized in our public schools. Furthermore, the fact that the psychological testing movement has been advancing with such great strides during the past fifteen or twenty years proves, to a certain degree, that it is filling that need.

Psychological tests have been accepted by leading educators as valuable aids to administration. This fact is apparent from the results of a questionnaire sent to the superintendents of public instruction and to directors of psychological testing in the school systems of all the cities in the United States with 100,000 or more population during the school year of 1928–1929. The questionnaire was returned from 86 cities, each one having all or part of the questions answered.

PURPOSE OF THE STUDY

The purpose of this study was to determine the status of psychological testing in the larger city school systems of the United States and to trace its developmental history. No attempt was made to justify psychological tests. "Mental testing requires no justification; the rapid growth of belief in its value has in many ways proved to be well warranted."[1]

[1] A Manual of Individual Mental Tests and Testing. By Bronner, Healy, Lowe, and Shimberg, p. 3.

DEVELOPMENTAL HISTORY

Fifty-four answers were received to the following question: "About what year were psychological tests first used in your school system?" The results are shown in table 1.

The earliest date reported was 1898 from Chicago. No information was given as to the type of tests used nor to what extent they were used. The next city to introduce testing was Boston in 1907 or 1908. Dr. Arthur C. Jelley of that city was given credit for being one of the first to apply psychological tests in the United States. In 1909 they were introduced into the

TABLE 1

Showing when psychological tests were first used in 54 cities of the United States

DATE	POPULATION				TOTAL
	100,000 to 250,000	250,000 to 500,000	500,000 to 1,000,000	Over 1,000,000	
1898				1	1
1901–05					0
1906–10	1	3		1	5
1911–15	7	3	4	1	15
1916–20	13	4	2	1	20
1921–25	9		2		11
1926–27	2				2

Rochester Schools and in 1910 into the schools of Indianapolis, Seattle, and Dayton. The latest date reported was 1927.

Tests were introduced into 53 out of the 54 cities that answered this question during the period from 1907 to 1927 inclusive. Twenty cities introduced the tests during the five year period from 1916 to 1920. This was the period of greatest development. The smaller cities, e.g., from 100,000 to 250,000 population, were slower to introduce the tests than the larger cities.

PSYCHOLOGY DEPARTMENTS

Psychological testing has become so important and has attained such large proportions that many cities have felt the need

of organizing separate departments to properly handle this work.

Twenty-one cities, about one-fourth of the total number which returned the questionnaire, reported separate departments. There are no data to show how well these departments are organized. Several cities may have reported separate departments when in reality they could not be classed as such. However, there is shown a strong tendency in the direction of departmentalizing psychological service.

The status of the psychology departments in four cities is questionable. Tacoma accompanied the affirmative answer with "Educational Research." This indicates that it is a part

TABLE 2

Dates of organization of psychology departments in 21 cities reporting separate departments

DATE	NUMBER OF CITIES
1899	1
1901–05	0
1906–10	0
1911–15	2
1916–20	10
1921–25	6
1926–28	2

of the Research Department. Jacksonville indicated that it is a division of the Department of Classification. The answer from St. Paul was, "Practically yes, technically no." The Research Department was also mentioned. Newark reported a separate department which is a sub-department of the Department of Child Guidance. However, since each of these four cities gave the affirmative answer they are listed with those having separate departments.

Chicago reported the first separate department which was organized in 1899. Seattle and Pittsburgh organized departments in 1912 and 1915 respectively. Ten cities organized separate departments during the period from 1916 to 1920.

(See table 2.) This was the period of greatest development of psychology departments.

Table 3 shows the departments supervising the psychological testing in the cities that have not yet organized separate depart-

TABLE 3

Departments in city school systems which supervise the work of psychological testing

DEPARTMENTS RENDERING PSYCHOLOGICAL SERVICE	POPULATION OF CITIES				TOTAL NUMBER OF CITIES
	100,000 to 250,000	250,000 to 500,000	500,000 to 1,000,000	Over 1,000,000	
Separate Departments of Psychology	10	4	4	3	21
Department of Educational Research and Statistics	14	7	5	1	27
Department of Special Classes, Special Education, etc	5	5	2		12
Mental Hygiene Clinic, Child Guidance Clinic, etc	1	1	1		3
State Clinic	1				1
Division of Tests and Measurements	1				1
Department of Health Service	1				1
Department of Testing and Guidance	1				1
Department of Records, Reference and Guidance	1				1
Bureau of Vocational Research	1				1
Department of Instruction	1				1
Department of Attendance, Social Service	1				1
By Members of Staff—Teachers, Principals. No department reported	6				6
By School Psychologist. No department reported	1				1

ments. The lack of uniformity of these departments is very noticeable, especially in the cities below 250,000 population.

The Research Department was reported most frequently (by 27 cities) as being responsible for rendering psychological service. Twelve cities reported the Department of Special Classes or

Special Education. Six cities reported that the testing work is done by members of the staff, e.g., teachers and principals. These cities did not mention any particular department. One city reported that the work was done by the school psychologist without mentioning any department. The other departments are shown in table 3.

Manhattan Borough, New York, did not report a separate department. Her psychological work is centered in the Bureau of Reference, Research, and Statistics, and the Department of Ungraded Classes. She has certain schools that provide facilities for giving vocational and educational advice to pupils based on the results of group and individual tests.

RANK OF PSYCHOLOGY DEPARTMENT HEADS

The following question was asked: "Does the Psychology Department Head have equal rank with other department heads?" Eighteen of the 21 cities with Psychology Departments answered, "Yes." One city answered, "Not Yet," and another answered, "Not in salary." The other city did not answer this question.

DEGREES HELD BY HEADS OF PSYCHOLOGY DEPARTMENTS

The seriousness with which psychological testing is being accepted is indicated by the fact that advanced degrees are held by the heads of most of the departments. The Ph.D. degree is held by nine of the 21 department heads, the M.A. by seven, the M.S. by one, and the B.S. by one. Three cities did not answer this question.

FULL AND PART TIME EXAMINERS

This study revealed the fact that psychological testing is being carried on almost universally in the larger cities of the United States. This is probably true of the smaller cities too. Only one city included in this survey reported that psychological tests are not being used.

Most of the cities which do not have separate departments employ full or part time examiners, or both. But, as a rule,

the cities which have separate departments employ many more full time examiners than those which do not, in proportion to their population.

Several of the cities utilize their teachers and principals for examiners. One city reported as many as 80 teachers devoting a part of their time to examining.

Table 4 shows the average number of full time examiners employed in cities with separate Psychology Departments as compared with the average number employed in cities without such departments.

TABLE 4

POPULATION	AVERAGE NUMBER OF FULL TIME EXAMINERS	
	Cities with Psychological Departments not including department heads	Cities not having Psychological Departments
Over 1,000,000	12+	
500,000 to 1,000,000	5	1+
250,000 to 500,000	3*	1+
100,000 to 250,000	1	1−

* Rochester belongs in this group but was omitted because of the exceptional number of full time examiners, a total of 14. This number would not change the median, but would make the mean much larger.

SALARY RECEIVED BY EXAMINERS

The general practice is to place the examiners on the same salary schedule as the teachers with equivalent training. Some cities stated that the examiners receive the same salary as high school teachers.

The salary of examiners was reported by 43 cities. The reports are summarized as follows:

Same salary as teachers	29
More salary than teachers	9
Less salary than teachers	5

The salary schedule is about the same in cities with and without separate departments. In both groups of cities the

salary of examiners is generally determined by the salary paid the teachers.

TRAINING REQUIRED OF PSYCHOLOGICAL EXAMINERS

Most cities require special training in psychological testing. Some cities require more special training than others. Eleven of the 35 cities answering this question require teaching experience in addition to the qualifications listed below. Two cities have examiners with Ph.D. degrees, but the highest degree required in each of these cities is the M.A. in psychology.

The qualifications required by 35 cities are summarized as follows:

Masters degree in psychology	11
Bachelor degree, psychology major	7
Special training in psychological testing	9
Selected on basis of competitive merit	1
No definite qualifications	7

Some of the cities did not mention degrees, but it is probable that they all require at least the bachelor degree.

ANNUAL COST OF OPERATING PSYCHOLOGY DEPARTMENTS

Fifteen cities with separate departments reported the annual cost of operating these departments. They are as follows:

Detroit	$75,000	Newark	$13,000
Rochester	52,500	Birmingham	12,000
Philadelphia	35,000	St. Paul	10,000
Buffalo	35,000	Toledo	10,000
Chicago	34,000	Kansas City	4,100*
Los Angeles	28,987	Fort Worth	3,750
Cleveland	25,392	Reading	3,200
Seattle	13,500		

* Kansas City's report does not include the salaries of the department head and assistants.

It is not certain that the above figures represent the total cost of psychological service in each city. In some cities certain

types of tests are given in other departments and in individual schools, the cost of which is not charged to the Psychology Department. However, a fair estimate of the approximate cost may be gained from these figures.

TYPES OF TESTS USED IN PUBLIC SCHOOLS

Table 5 shows seven types of tests being used together with the number and percentage of cities with and without separate departments using each type. A total of 69 cities reported the tests they are using.

TABLE 5

Types of psychological tests being used in public schools

TYPES OF TESTS BEING USED IN PUBLIC SCHOOLS	21 CITIES WITH SEPARATE DEPARTMENTS		48 CITIES WITHOUT SEPARATE DEPARTMENTS	
	Number of cities	Per cent	Number of cities	Per cent
Individual tests of general intelligence	21	100	43	92
Group intelligence tests	19	90	45	94
Performance tests	18	86	26	54
Special aptitude tests	16	76	21	44
Educational tests	21	100	42	88
Vocational guidance tests	11	52	11	23
Anthropometric measurements	10	47	3	6

It is shown in table 5 that the cities with psychology departments use a wider range of tests than cities without. Group intelligence tests, however, are used more extensively in cities not having separate departments. All of the other types of tests are used more extensively in cities with departments.

The Stanford-Binet was reported by 100 per cent of the cities with departments and by 90 per cent of the cities without. One city without a separate department reported using the Kuhlmann-Binet instead of the Stanford-Binet. Several cities without departments use only group tests for measuring general intelligence.

USES BEING MADE OF PSYCHOLOGICAL TESTS

A common criticism of psychological testing has been that after much time and money have been expended for purchasing, giving and scoring the tests, nothing is ever done about it. There are no data at hand to show to what extent this condition is true today, if indeed it is true. On the other hand, this study has revealed many uses that are being made of these tests in the larger cities. These findings should be suggestive to those cities which may not yet have discovered their use.

TABLE 6

Uses being made of psychological tests in 21 cities with and 48 cities without separate Psychology Departments

USES BEING MADE OF PSYCHOLOGICAL TESTS IN 69 CITIES	21 CITIES WITH SEPARATE DEPARTMENTS		48 CITIES WITHOUT SEPARATE DEPARTMENTS	
	Number	Per cent	Number	Per cent
Sectioning classes throughout entire system..	16	76	20	42
Sectioning classes in just a few schools.....	4	19	25	52
Determining low mentality for exclusion from school.....	16	76	36	75
Basis for vocational guidance.....	12	57	22	46
Recommendation for work permits.....	11	52	17	35
Aiding adjustment of problem children.....	19	90	42	88

Six general uses were listed in the questionnaire with instructions to check each one being utilized in the school system. Table 6 shows the results. Additional uses were also asked for. Sixty-nine cities responded to this section of the questionnaire including the 21 cities with separate departments.

This table shows that the general sectioning of classes throughout the entire school system is more prevalent in cities with separate departments than in cities without. The percentages are 76 and 42 respectively. However, in addition to those percentages, 19 per cent of the former and 52 per cent of the latter cities section classes in some of the schools. Only one of the former and three of the latter cities do not section classes

at all. Thus, 94 per cent of the 69 cities section classes in at least some of the schools by means of psychological tests. This is the most general use being made of the tests as judged by results of the questionnaire.

The second most common use made of the tests is, "*Aiding in the adjustment of problem children.*" Ninety per cent of the cities with departments and 88 per cent of the cities without reported this use.

Many additional uses were reported by these cities. They are as follows:

Selection of pupils for atypical classes
For high school graduation
To determine eligibility to high school of non-resident pupils
To classify beginners
For entrance into first grade
For promotion
To aid in supervision
For appraisal of curriculum
To evaluate methods of teaching
To objectify standards of achievement
For scholarship candidates
For candidates to skip one-half grade
To determine promotion of doubtful cases
To select cases for psychiatric and neurological observation
For study of juvenile court cases
For study of cases referred by social agencies

SUMMARY OF FINDINGS

1. The first city to use psychological tests was Chicago in 1898.

2. Fifty-three out of 54 cities which answered this questionnaire introduced psychological tests into their schools between 1907 and 1927 inclusive.

3. The five year period of greatest development of testing was from 1916 to 1920, during which period tests were introduced into 20 cities.

4. Twenty-one cities reported separate Psychology Departments. They comprise about one-fourth of all the cities responding.

5. The first separate Psychology Department to be organized in any school covered by this survey was in Chicago in 1899.

6. During the five-year period from 1916 to 1920 separate departments were organized in 10 cities. This was the period of greatest development of Psychology Departments.

7. In cities not having separate departments the Research Department is most frequently responsible for psychological testing service.

8. Eighteen of the 21 Psychology Department heads rank equally with the heads of other departments in their respective school systems.

9. The Ph.D. degree is held by nine of the 21 Psychology Department heads, the M.A. by seven, the M.S. by one, and the B.S. by one.

10. The cities having separate departments have many more full time psychological examiners than the cities not having separate departments.

11. The common practice is to pay the psychological examiners the same salaries received by teachers with equivalent training.

12. Most cities require special training in psychology to qualify psychological testers. Many cities require a degree in psychology. Some cities also require teaching experience.

13. Most types of psychological tests are being used by school psychologists. The Stanford-Binet is used almost universally.

14. Cities with Psychology Departments use a wider range of tests than cities without such departments.

15. Psychological tests are used to section classes in 94 per cent of the cities answering this section of the questionnaire.

THE BLOOD PRESSURE AND GALVANIC REFLEX AS INDICATORS OF EMOTIONAL STATES

L. E. RACKLEY

University of North Carolina

There has been considerable work done with the galvanometer measuring emotional states produced in different controlled conditions. All the investigators seem to agree that the galvanic reflex indicates the emotional state of the individual. However, there are different theories as to just what the reflex depends on. Since it is not the purpose of this paper to enter into any discussion in favor of any of these theories the interested reader is referred to the wealth of literature on the subject.[1]

There have been some investigations carried on to discover the relation of blood pressure change to emotion. The problem was investigated in 1895 by Binet and Verschide.[2] This investigation was carried out on four subjects and the experimenters recorded blood pressure change of five mm. hg. after fear-producing stimuli had been applied. In 1917 Marston[3] carried on an investigation using ten subjects to determine the change in blood pressure produced by conscious deception. And in 1923 he repeated this experiment on other subjects using both male and female to determine the sex difference in blood pressure change produced by conscious deception. Each S, after choosing to tell a lie or the truth about a certain matter was placed on the witness stand and cross examined about the statements he had made. During this time the blood pressure

[1] A very good description of the galvanic reflex from the standpoint of physiology and neurology, and also of the galvanometer, is given by Wechler in the Archives of Psychol., 1925, **15**.

[2] L'Annee Psychologique, 1897, **3**, p. 125.

[3] J. Exper. Psychol., 1917, **2**, p. 117; Ibid., 1923, **6**, p. 367.

was taken by the experimenter after a certain number of questions had been asked. Then he judged whether or not the S was telling a lie by the blood pressure results. He recorded an average blood pressure change of 12.6 mm. in cases where S chose to tell a lie, while he recorded as the highest blood pressure rise in cases where S chose to tell the truth, 6 mm. He also recorded more blood pressure change for the women than for the men. In 1923 Larson[4] carried out an experiment in the school for policemen in Berkeley, California. He used prisoners as S's. Each S was placed on the witness stand and was questioned about the crime which he was accused of committing. During this examination the experimenter took the blood pressure at regular intervals noting especially the blood pressure change produced by certain questions. The experimenter used the blood pressure change as a basis of determining the guilt or innocence of the prisoner and checked his results by the later confessions of the subjects. He recorded blood pressure changes as great as 26 mm. in case of one criminal and 30 mm. in another. In 1925 Landis and Gullette[5] in the University of Minnesota carried on an experiment using twenty-five subjects to determine blood pressure and inspiration-expiration ratio. They used as stimuli deception (S was instructed to lie one time and tell the truth one time), sudden noise (bursting of fire cracker near S), mental work (multiplication, 79×67). They recorded average blood pressure change as high as 3.2 mm.

So far as the writer knows no scientific work has been done to determine the amount of blood pressure change necessary to indicate the presence of emotional excitement in the S. However, Marston[6] suggests that there should be blood pressure change of 8 mm. or more to be significant.

The writer became interested in this problem while working as a graduate student in the Jesup Psychological Laboratory at Nashville, Tennessee, under the direction of Dr. Joseph Peterson in 1927 and 1928. A preliminary study was carried out

[4] J. Exper. Psychol., 1923, **6**, p. 420.
[5] Psychol. Monograph, 1924, **33**, No. 150
[6] J. Comp. Psychol., 1925, **5**, p. 221.

in the laboratory on about forty-five students in psychology during the fall quarter, 1927.

S	Part	First reading		Second reading		Third reading		Fourth reading	
		s	d	s	d	s	d	s	d
1	O	91.75	65.50	91.75	64.50	93.50	65.75	94.25	68.00
	E_1	80.00	56.00	86.00	64.00	89.00	66.00	89.00	70.00
	E_2	108.00	72.00	122.00	84.00	126.00	84.00	126.00	88.00
5	O	100.25	64.75	97.00	64.75	99.00	64.75	99.25	64.25
	E_1	96.00	66.00	103.00	74.00	97.00	70.00	100.00	70.00
	E_2	90.00	62.00	108.00	80.00	106.00	78.00	110.00	80.00

b p

S	Part	Av. b p		Av. Deviations	
		s	d	s	d
1	O	92.81	65.94	1.59	1.59
	E_1	86.00	64.00	3.00	2.50
	E_2	120.50	82.00	6.25	5.00
5	O	94.49	64.61	1.07	.66
	E_1	99.00	70.00	2.50	2.00
	E_2	103.50	75.00	6.75	6.50

PE_{DIF}

S	Parts	D		PE_{DIF}		$\frac{D}{PE_{DIF}}$		D in sigma units	
		s	d	s	d	s	d	s	d
1	OE_1	6.810	1.940	1.435	1.253	4.746	1.548	3.417	.974
	OE_2	27.690	16.060	2.726	2.217	10.121	7.244	3.535	2.563
	E_1E_2	34.500	18.000	2.931	2.456	11.737	7.329	4.404	5.745
5	OE_1	.490	5.390	1.392	.890	.352	6.056	.365	2.150
	OE_2	3.010	10.390	2.993	2.761	1.006	3.763	.356	1.275
	E_1E_2	4.500	5.000	3.043	2.874	1.479	1.736	.532	.614

b p

S	Part	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16*
1	O	.146	.231	.173	.163	.151	.156	.155	.166	.170	.160	.144	.145	.145	.139	.141	.143
	E_1	.160	.725	.750	.725	.635	.520	.525	.550	.560	.475	.440	.550	.540	.575	.605	.605
	E_2	.090	.175	.175	.135	.150	.175	.180	.225	.255	.250	.240	.240	.300	.255	.240	.215
5	O	.155	.149	.146	.148	.138	.138	.146	.156	.103	.155	.151	.146	.146	.143	.143	.151
	E_1	.200	.220	.200	.195	.175	.140	.155	.225	.140	.160	.130	.185	.170	.180	.165	.175
	E_2	.250	.090	.090	.085	.070	.065	.065	.065	.060	.105	.085	.085	.090	.075	.095	.075

Av. g r's

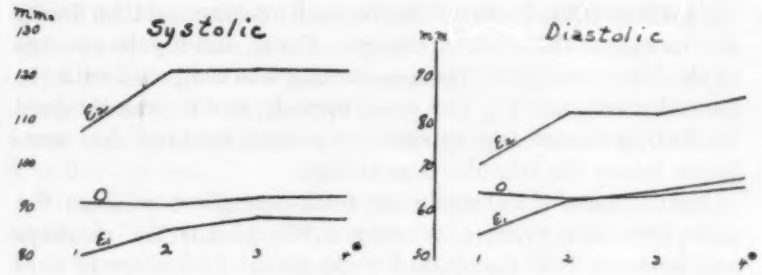
* Successive readings.

It was the purpose of the present experiment to determine the effect of fear-producing stimuli and of rapid mental work on blood pressure and the galvanic reflex. It was carried out at the Masonic Home at Nashville, Tennessee, during the summer quarter, 1928. Ten subjects were used, five boys and five girls, ranging in age from thirteen to seventeen. Four of them were in high school and the others were in the grades. The blood pressure was taken with a Baumanometer and a Bowles Stethoscope and the galvanic reading with the Hathaway Psycho-galvanometer.

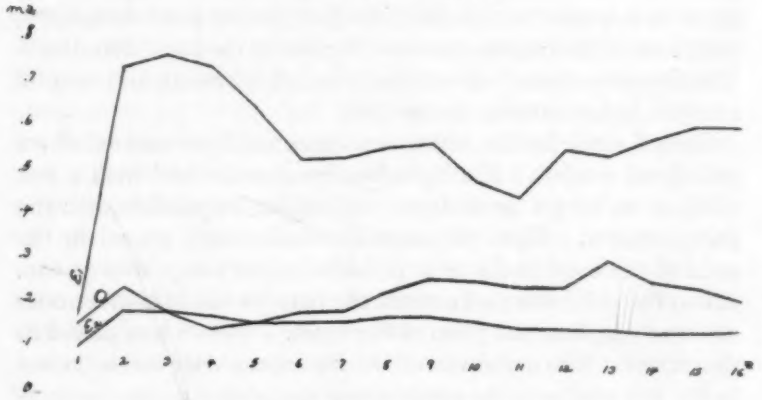
The following abbreviations of terms will be used: b p for blood pressure, g r for galvanic reading, O for part one of the test when no stimulus was used, E_1 for part two of the test when the substitution test was used as a stimulus, E_2 for part three of the test when the fear-producing stimulus was used, differences will be abbreviated as D. The b p readings are all given in millimeters and will be abbreviated as mm., all g r are given in milliamperes and will be abbreviated as ma. Systolic will be represented by (s), and Diastolic by (d).

PROCEDURE

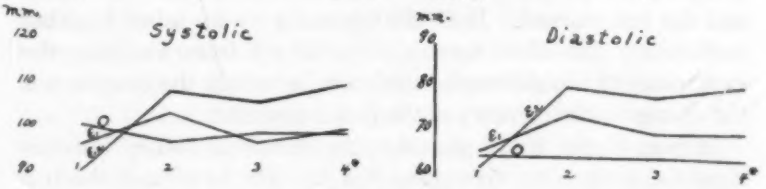
In the whole study the b p was taken four times each evening at intervals of two minutes, and the g r was taken every fifteen seconds for eight minutes. The tests for part O lasted four evenings; those for parts E_1 and E_2 lasted for one evening only on each S. Since it was the plan of the study to compare the normal reactions of the S's with the reactions of the same S under different stimulation, the normal reading was found by getting



S 1 - B.P.



S 1 - G.R.



S 5 - B.P.

S 5 - G.R.

the average of the four readings for each evening and then finally the average of those four evenings. For E_1 and E_2, the average of the four readings for the one evening was compared with the normal reaction. The first g r in parts E_1 and E_2 was obtained by finding the average of the four normal readings that were taken before the stimulus was applied.

For all parts of the study the tests were given at about the same time each evening, between 6:30 and 8:00 for the boys and between 7:30 and 9:00 for the girls. The subjects were tested in the same order each evening, those of the girls being given in a spare room on the third floor in the girls' dormitory, and those of the boys in the writer's room in the boys' dormitory. The S's were asked to refrain from all physical and mental exertion before coming to the test.

Each S came for the test at an appointed time and sat down quietly at a table. His right hand was moistened with a wet cloth so as to get as uniform conduction as possible with the galvanometer. Then the large electrodes were placed in the palm of the hand and a piece of folded paper was placed on top. A handkerchief was tied around the hand to hold the electrodes uniformly against the palm of the hand. Then S was placed in the circuit. The arm-band of the Baumonometer was adjusted to the left arm[7] and the stethoscope was placed under the arm-band in the proper position. The time was carefully noted and the test started. Both the b p and g r were taken together each time. No effort was made to keep S from watching the excursions of the galvanometer while he was in the circuit, nor the change in the mercury of the Baumanometer.

In part E_1 the S was placed in the circuit as before, but this time the electrodes were placed in the left hand and the b p was taken on the left arm as usual. This was necessary in order to keep the right arm free for writing. The substitution test was placed face down on the table before S with a pencil beside it. After explaining to S just what he was to do when the signal was given, the b p was taken once and the g r was re-

[7] This was reversed in case of subject 7 who had a boil on her left arm just above her elbow.

corded at intervals of 30 seconds for two minutes. Then he was told to turn the paper over and do the substitution as fast as he could until he was told to stop. The b p was taken three times while S did the substitution, the g r being recorded as usual. At the end of six minutes the experimenter said "time" and the test ended.

In part E_2 the b p was taken once and the g r was recorded at intervals of 30 seconds for two minutes. Then at a signal, Mr. L. C. Connell, Superintendent of the Masonic Home, apparently without pre-arrangement came into the room and started talking to S. As he talked to S both the b p and g r were recorded as usual. For some of the S's the conversation was a correction for something which they had done, for others it was a questioning about some act which they were suspected of doing, for others it was concerning future plans or other topics of vital interest to S. It should be emphasized here that all of these situations were real. S number 1, a 13 year old boy, was corrected for not returning from a visit to his people when he had promised to return. S number 2, a 14 year old boy, was corrected for not looking after his paper route in the proper manner and he was told that he would probably lose it. S 3, a 13 year old boy who has a brother in the reform school for stealing, and who himself has been guilty of stealing, was questioned about stealing. S 4, a 13 year old boy was questioned about disorderly conduct in the dormitory. S 5, a 16 year old boy, drove a truck for the home, a job which he liked very much. He was questioned about speeding the truck, which he admitted he had done, and was told that if it happened again he would be taken off the truck. S 6, a 15 year old girl, was condemned for not practicing her music faithfully. S's 7 and 9, 15 year old girls, were questioned about future plans and told that they would soon have to leave the home and provide for themselves. S 8, a 15 year old girl, had broken the rules of the institution and had told the superintendent a lie about it. At this time she was questioned further about the affair. (This case is very much like the case given by Marston.[3]) S 10, a 17 year old

[3] Cf. foot note 3.

girl, was told some bad news about her home which she did not believe.

RESULTS

The readings obtained were at first higher than later on in the test. It would seem that this difference is due in part, perhaps, to the fact that the situation was new, and as soon as the S became acquainted with it, the tension seems to have been released, and the b p and g r are both somewhat lower. The results obtained in part O of the test vary in b p about four to seven mm. and the g r about .1 to .3 ma.; the results obtained in part E_1 vary in b p from 1 to 12 mm. (s), and from 4 to 14 mm. (d), and in g r from .08 to .35; the results of part three vary from 7 to 32 mm. (s) and from 10 to 24 mm. (d), and in g r from .04 to .325. Since no sex difference was found, the data are treated all together.

It should be kept in mind that it is the plan of this study to compare the results of the three parts of the test, rather than compare the results of one S with those of another. Typical results are presented in tables and graphs on pages 399 and 401.

CONCLUSIONS

The present study seems to warrant the following conclusions:

1. That b p changes and g r changes indicate the presence of emotional states.

2. That fear-producing stimuli cause a greater change in b p and g r than mental work.

3. There is a positive relationship between the b p and the g r results, but this relationship is not great.

NEWS AND NOTES

Dr. Donald A. Laird, Director of Colgate University Psychology Laboratory, has been appointed to write the sections on "Sleep" and "Fatigue" for the *Psychology of Medicine*, under the editorship of Dr. George M. Piersol, the post-graduate school of medicine at the University of Pennsylvania. Dr. Laird has also been appointed to prepare a chapter on "Work and Fatigue" for the *Oxford Medicine* under the editorship of Dr. Henry A. Christian, the Peter Bent Brigham Hospital, Boston.

The following announcements come from Stanford University: Mr. Albert Walton of the University of Washington is to be acting instructor of psychology at Stanford for the coming academic year. Dr. Keith Sward of the University of Minnesota who spent the year 1929–1930 at Stanford University as a Social Science Fellow has gone to Western Reserve University as assistant professor of psychology.

The National Safety Council, New York City, has just completed a statistical study of student accidents from September 1929 to June 1930 in which it was found that out of a total of about 29,000,000 pupils in grade and high schools there were 560,000 accidents, 5,300 of them causing death. The most outstanding fact is the predominance of home accidents, constituting nearly 32 per cent of the total. These are followed by "other" accidents which occur at places other than at home or at school and outside school hours, which total 24 per cent. The magnitude of the child accident problem certainly calls for intelligent action based on facts.

W. W. Charters, president of the Research Division of the Ohio Education Association, called a meeting of the members of the Division and other persons engaged in educational research throughout Ohio, on Saturday, October 4. The purpose of this meeting was to discuss the possibility of arranging a program of co-operative research to be sponsored by the Research Division and to be carried on by the members and other persons interested. A Committee on Co-operative Research Activities of the Research Division was appointed by the president with Dr. R. L. Morton, Ohio University, chairman. The duties of the committee were stipulated as follows: to prepare a program of co-operative research and arrange the preliminary procedure which will be discussed at the December meeting of the Research Division; to set up the speci-

fications for these research projects; to endeavor to coördinate all the research agencies in the state; and to call special meetings to discuss the work as it progresses.

A central Ohio conference on supervision and teacher training will meet at Ohio Wesleyan University, Delaware, Ohio, November 21 and 22, 1930. General sessions will be devoted to problems of subject-matter courses for teachers, work of critic teachers, practical problems of supervision, and applied work in supervision. There will be several sectional meetings. Dr. Francis D. Curtis, University of Michigan, will appear on the sectional program of the natural science group.

The Board of Governors of Mooseheart, Illinois, the "City of Childhood," established by the Loyal Order of Moose for orphaned children of its membership, announces through its Superintendent, Mr. Ernest N. Roselle, the establishment of a new Department of Child Research.

Dr. Martin L. Reymert, formerly head of the Department of Psychology at Wittenberg College, has been appointed Director of Child Research, and is now establishing a laboratory at Mooseheart devoted to all phases of child research.

BOOK REVIEWS

MURCHISON, CARL, ed. *A History of Psychology in Autobiography.* Vol. I. Clark University Press, 1930, pp. xvii + 516.

The value of the historical approach to a broad foundation in any science can hardly be questioned. Biography affords an important source in the study of history and for the young student autobiography gives a personal contact with the lives of men not to be found in any other sort of written record. This is especially true if the author is generous in including those "certain aspects of scientific work that usually find no formal expression, such as personal background and bias, the methods of constructive thinking, and the appraisal of problems" (Dodge).

The volume being reviewed is the first of a projected series of short autobiographical sketches in which the authors are asked to relate "as much of their personal history as bears on their professional careers." The plan cannot be commended too highly; the fruition is directly dependent on the authors.

Psychologists now at work in England, France, Germany, Italy, Switzerland and the United States, and one physiologist from Holland are included in this volume. Assuming that the requests made to each person were uniform it is interesting to note the different reactions—personal not national.

Claparède, Dodge, Jastrow, McDougall, Seashore and Warren have painted delightful pictures of their early years and of later activities. Their reminiscences make them seem human rather than academicians or austere scientists. I mention these men together because they seem to me to come closest to the program of these volumes as I understand it.

The others, while equally interesting, have paid varying degrees of attention to their philosophic or psychologic systems and in so far as they do they lose the personal value. The intellectual life of a man can be interpreted largely from his writings so that a mere resumé of ones bibliographical contributions seems somewhat out of place. Likewise a defense of one's systematic position can hardly be classed as autobiography. Without comment I shall list the remaining contributors—Baldwin, Calkins, Janet, Kiesow, Spearman, Stern, Stumpf, Ziehen and Zwaardemaker. Many of these have written biographical material of great interest—others have not, but I must leave it to the readers of the book to separate the sheep from the goats.

Inasmuch as one of the purposes of this series is to afford younger men

some examples of the professional development of the older workers it might be interesting to attempt to analyze the contributions to find, if possible, one or more uniform circumstances to which these people attribute, directly or implicitly, their success.

The academic training and youthful aspirations vary. Many were medical men, a few had ministerial inclinations, some started in philosophy, the minority had an early interest in psychology. Financial status varied widely as did family position and, later, academic position and advancement. None of these things present any degree of uniformity, but there are two facts which the present reviewer finds universal in this group.

The first of these universal conditions sounds platitudinous—but like most platitudes has a basis in fact. It is characteristic of these men that they *worked*. Not work of the sort that is imposed by extraneous influences, but work as the result of a deep-seated desire to know and to accomplish. The second condition may not be so important but it is notable that each of these persons is able to mention definitely at least one older man who had a direct and beneficial influence on his life. In perhaps the majority of cases this advisor was a teacher, often it was a relative or friend. But I must not labor the point; it is hardly the task of a reviewer to produce an inspirational essay. I present these two major facts and urge all the younger psychologists to read the volume for the details.

It is not possible to close this review without one word on the typography. Minor errors such as incomplete parentheses, poor spacing, etc., are prevalent, but there are three glaring errors to which attention must be called. On page 155 there is a completely inverted line. In the first paragraph on page 248 one or more lines have been omitted resulting in the complete loss of a pertinent point. The third instance is the use of the figure "*1*" instead of the italic "*l*" in the formulae on pages 510 and 512. Surely a work of this character deserves more careful proofreading!

C. M. Louttit,
Ohio University.

Valentine, Willard Lee. *Readings in Experimental Psychology.* Columbus, Ohio, R. G. Adams & Co., 1930, pp. 146.

These readings consist of 27 abstracts of 24 articles, four of these coming from parts of Clark L. Hull's article, "Quantitative Aspects of the Evolution of Concepts." Four abstracts are on the general subject of reaction times and a fifth, "The Accuracy of Timing with a Stop Watch," is very closely related. There are two studies on hunger behavior, one on sensori-motor responses in infants and two on emotional behavior, one of these being Mandel Sherman's study of emotional

responses in infants. There are two articles on delayed responses, one on sleep, one on the relative importance of accommodation and convergence in distance perception and one on local signature. Under the general head of learning are found studies on mirror drawing, drives in animal learning, comparison of human beings and rats in maze learning, relation between learning and amount to be learned, multiple-choice method in the study of primates and the studies by Hull mentioned above. In addition there is a study of the correlation between intelligence tests and scholastic success, and a 20-page article on sensory discrimination.

These readings are intended as "parallel" for an introductory course in psychology, and seem to the writer to be rather well chosen. Most of them cover pertinent points and represent the latest findings. Were the writer to make any change in this collection it would be largely one of addition with one or two possible substitutions.

This collection, according to the author's statement in the preface, "fulfills two needs of the elementary student. In the first place, it introduces him to experiments which are treated sketchily and secondhandly in the formal text." Library facilities, as a rule are too inadequate for the whole class to read the original articles and when assignments are made for class reports, quite frequently the student making the report is the only one who gets anything out of it. "This sourcebook is designed to remedy this situation by putting a copy of the material in almost its original form in the hands of every student."

The writer believes that the author is somewhat optimistic in the extent to which he believes this need is fulfilled. These abstracts are an improvement on the sketchy treatment in the text, but the articles can hardly be said to be "almost" in their original form, since they have been abbreviated from a third to a tenth of their original size. The writer also doubts whether the elementary student has quite the respect for an abstract that he has for a "selection," even though the abstract may be more inclusive, and also at times more to the point. The articles included in this collection have lost none of their original clearness and in some cases it has been improved.

"The second need," according to the author, "concerns courses in which laboratory is offered. For most students the collection of data presents no problem, but the analysis and presentation of it in acceptable fashion is a serious stumbling block. In practically all of the experiments included in this compilation I have made an effort to present "problem,' 'data,' 'discussion' and 'results' in an acceptable report style. Thus they are to serve as models in the preparation of an original report from data collected by the student in the laboratory."

This need, it seems, is adequately fulfilled by these readings.

F. C. Scott,
Ohio University.

GALLAGHER, RALPH P. *Courses and Careers.* Harpers & Brothers, 1930, $1.40, 404 p. Guidance Adviser, Elizabeth Public Schools, Elizabeth, New Jersey; Guidance Instructor, Tulane University Summer School.

It is of interest in examining a new book in a given field to make a comparison with the more or less standard treatment of the subject. The purpose of a text book in vocational guidance is that of informing about occupations and of social training through giving an understanding of the social contribution of the work of other groups of people. For successful participation in occupational activities certain moral and social elements must be considered as well as a study of economic factors.

Gallagher's *Courses and Careers* aims to present the materials needed through lesson plans organized in the way familiar to both teachers and pupils of (1) a list of approved references to which the pupil must go before treatment of the lesson, (2) a plan of note-book work, and (3) a list of some twenty questions.

The reference list is given so much importance in the lesson that the book itself might be described as a manual for study accompanying any text book. For occupational information, then, the reference readings make up for inadequacies of the book. The guidance note-book plan is carried out in every lesson consisting of copy-material and problem-solving, a monotonous procedure unless activated by extra assignment for outside investigation. For purposeful activity the book looks to the teacher for inspiration. Providing each pupil used his own text it would be a waste of effort and time to keep a note-book copy. The twenty or more questions put up for discussion at the end of the lesson are valuable in so far as they are satisfactorily answered. Some of them seem ambiguous if at all answerable. A class might well go astray in its consideration of them without the competent leadership of a well informed and skillful teacher.

As a guide for organizing a class for the study of occupations it meets the problems of the teacher new in the field of guidance, and will furnish a basis for departure in planning the program for the course. It is recommended for Junior and Senior high school work in vocational guidance.

CORNELIA HOLCOMBE,
Ohio University.

G. M. STRATTON. *Social Psychology of International Conduct.* D. Appleton & Company, New York, 1929. x + 387 pp.

Dr. Stratton's book belies the not infrequent view that social psychology can contribute little toward control of the forces—mainly war—which at times threaten civilization. Here is a closely packed summary

of the matured wisdom of a relatively unbiased student of international conduct, presented in an easy persuasive style that has clarity and is usually convincing. The volume will be found very acceptable to those who have sought for a text on the psychology of international relations or for a second semester course in advanced social psychology. Because of the non-technical treatment the lay reader will find it easily within his powers of comprehension.

In the present confusion of ideas about such controversial subjects as "unchanging human nature," international coöperation, crowd-mind behavior, backward peoples and imperialism, national self-determination, racial attitudes and kindred topics, Dr. Stratton's treatment will be welcomed. Many of the theses put forth by the author are unlike those we have known; some are in distinct disagreement to some ancient theories: but all of his views are abundantly supported by reference to actual international incidents or familiar situations.

The volume is organized in three parts. Part I, Preparation for International Behavior, is a discussion of racial evaluations, in which experimental, historical, and statistical findings are marshalled forth to explain the mental likenesses and contrasts in racial stocks. The conclusions are not essentially different in their broad outlines from other treatments but are carefully drawn. All races have backward stocks yet the proportion of backward to advanced strains separates the stocks widely. The greater proportions of advanced strains in the Mongoloid and Caucasoid races are attributed to a superior psychic endowment favorable to organization and perpetuation through social inheritance. The reviewer senses an effort to examine the evidence in a sincere and sympathetic manner, yet it must be said that few fields offer more difficult material from which conclusive findings can be wrested than from racial comparisons. Yet here is distinct progress. At least, there is no scientific offering to the Nordic god! Prejudice is found to be widely disseminated both in time and geographically, yet such universality does not lead to the gratuitous assumption of innateness. Rather he lays it to the door of social and economic expediency, showing that it is roughly proportional to the need of defense from threatened injury. By much the same reasoning it is shown that nationality, national spirit, patriotism, resolve themselves more or less into fixed attitudes within the individual, subject to the waxing and waning to which all other attitudes are subject. And in this conception Stratton introduces the view that the national "mind" like any individual mind, may have its delusions, its obsessions, its depressions and its elations, its conflicts and its logic-tight compartments:

Part II, The Conduct of Nations toward One Another, presents material that has only infrequently crept into scientific literature. Prior to Allport, students of group psychology have been inclined to neglect

the individual; specifically, that the individual's feelings have risen and fallen with the fate of the nation. In the Stratton view the internationally-significant affairs of the state have usually been a concern first of the particular individuals involved in activities having international complications (the active-interests group). These private and special objectives are then made the general concern of the whole group by the invoking of popular motives. Desires thus drive the international life, and their spatial distribution determines the pattern and reconciliation of *national* desires. Thus the consistent friendliness of United States and Brazil can be explained as well as the recurrent repulsion between Germany and France. The sections devoted to conflict behavior disclose fresh viewpoints. In striking contrast to the current and persistent notion that man's unchanging nature makes wars forever inevitable, the author shows by analogy to the one-time accepted institutions of blood revenge and slavery, that riddance of war requires no change in human nature. War, it is argued with much evidence, is a product of social planning. It is not a relapse into cave-man status; not an ebullition of surging and recurrent pugnacity. It is a product of governmental art; not something which the individual in blind unreason plunges. Indeed, war's sources are found in the *intelligence* of man: when war in general seems to come in for disapproval, defensive wars are sanctioned, notwithstanding the observation that to the nation itself most wars are defensive. Against even the hoary notion, so fondly held by McDougall, that the aggressively pugnacious animals and men are superior, Stratton marshalls evidence that the jungle or the mountain fastness illustrates the law that the animal lives longest, thrives best, who *avoids* fighting except when absolutely necessary.

The concluding section, being devoted to the outlook for inter-group coöperation, is of less interest to the psychologist, more to the general reader. Encouragement is suggested in the demonstrated educability of pugnacity, in the possibility of relief from the crowd psychologists. The book concludes with suggestions for reëducation in attitudes, particularly toward war. The reviewer feels that however praiseworthy the intention of the author, the adding of a social program for the advance of mankind detracts from the full power of the book as a pioneer work in scientific group-psychology.

NORMAN C. MEIER,
University of Iowa.

DAVIS, ROBERT A. *Mentality of Orphans.* Richard G. Badger, Publisher, Boston, 1930, 182 pp., $5.00.

The material on which this monograph is based was obtained by giving mental examinations to orphans in six Texas orphanages. Two group tests were given to each child. The results obtained were largely cor-

roboratory. The major results indicate (1) that orphans are a select group, being inferior to public school children in Texas, (2) that teachers in orphanages overestimate their pupils and their school attainments more than do teachers in the public schools, and (3) that life in orphanages has little effect upon the mentality of the child. This last point was based upon a study of siblings in the orphanages. The author comments upon social and educational problems dealing with orphans, in the light of his findings.

STUART M. STOKE
Ohio University.

NEW BOOKS AND PAMPHLETS RECEIVED*

Books and pamphlets for review should be sent to James P. Porter Department of Psychology, Ohio University, Athens, Ohio.

Academic Prognosis in the University. HAROLD A. EDGERTON. Warwick and York, Baltimore, 1930. $1.88. 83 pp.

The Behavior of the Newborn Infant. KARL CHAPMAN PRATT, AMALIE K. NELSON, AND NOW HUA SUN. Ohio State University Press, 1930. Contributions in Psychology, No. 10. Columbus, Ohio. 237 pp.

Civil War Prisons: A Study in War Psychology. WILLIAM BEST HESSELTINE. Ohio State University Press, Columbus, Ohio. 1930. 290 pp.

Conflicting Psychologies of Learning. BOYD HENRY BODE. D. C. Heath & Company, 1929. 305 pp.

Courses and Careers. RALPH P. GALLAGHER. *Harper & Brothers,* New York City. 1930. $1.40. 404 pp.

Educational Measurement in the Elementary Grades. I. N. MADSEN. World Book Company, Yonkers-on-Hudson, N. Y.

Foundations for Human Engineering. CHARLES R. GOW. Macmillan Company, New York, 1930 226 pp.

The Growth of Ability. RAYMOND O. FILTER AND OMAR C. HELD. Warwick and York, Baltimore, 1930. $2.28. 174 pp.

A History of Physiology in Autobiography, Vol. I. CARL MURCHISON, Editor, and fifteen authors. Clark University Press, Worcester, Mass. $5.00. 516 pp.

Introduction to the Psychology of the Classroom. CHARLES ELMER HOLLEY. D. C. Heath & Company, 1930. 257 pp.

Measurement of Man, The. J. ARTHUR HARRIS, CLARENCE M. JACKSON, DONALD G. PATERSON, AND RICHARD E. SCAMMON. University of Minnesota Press, 1930, Minneapolis, Minn. $2.50. 215 pp.

Mental Aspects of Stammering. C. S. BLUEMEL. Williams & Wilkins Company, Baltimore, Md. $2.50. 152 pp.

Mentality of Orphans. ROBERT A. DAVIS. Richard G. Badger, Publisher, Boston, Mass. $5.00. 182 pp.

Psychology for Advertisers. D. B. LUCAS AND C. E. BENSON. Harper & Brothers, Publishers, New York. 1930. $3.50. 351 pp.

---

* Mention here does not preclude further comment.

The Psychology of the Common Branches. WILLIAM HENRY PYLE. Warwick and York, Baltimore, 1930. $2.00. 381 pp.

Readings in Experimental Psychology. WILLARD LEE VALENTINE. R. G. Adams & Company, Columbus, Ohio. 146 pp.

The Rorschach Test and Personality Diagnosis. SAMUEL J. BECK, Reprinted from American Journal of Psychiatry, Vol. X, No. 1, July 1930. 52 pp.

Stammering. ELSIE FOGERTY. E. P. Dutton & Company, N. Y. C. 1930. 95 cents. 64 pp.

Strategy in Handling People. EWING T. WEBB AND JOHN J. B. MORGAN. Boulton, Pierce & Company, Chicago, 1930. $3.00. 260 pp.

What White People Can Do to Promote Negro Education. George Peabody College for Teachers. Vol. XIX, No. 8, October, 1930. Nashville, Tenn. 58 pp.

Directory of American Psychological Periodicals (*Continued*)

Psychological Abstracts—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter. Monthly. Abtracts of psychological literature. Founded 1927.

Journal of General Psychology—Worcester, Mass.: Clark University Press. Subscription $7. 500–600 pages annually. Edited by Carl Murchison. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Worcester, Mass.: Clark University Press. Subscription $7.00. 500–600 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.

Journal of Personnel Research—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages approximately. Edited by Walter V. Bingham, and Clarence S. Yoakem. Bi-Monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.; Cornell University. Subscription $6.50. 624 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Journal of Genetic Psychology and The Pedagogical Seminary—Worcester, Mass.: Clark University Press. Subscription $7. 600 pages annually. Edited by Carl Murchison. Quarterly. Child behavior, differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.50. 540 pages annually. Bi-monthly. General Psychology. Edited by Howard C. Warren. Founded 1894.

Psychological Monographs—Princeton, N. J.: Psychological Review Company. Subscription $3.00 per vol. 500 pp. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.: Psychological Review Company. Subscription $4.00. 300–400 pp. Edited by W. S. Hunter. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Edited by Edward S. Robinson. Founded 1904.

Archives of Psychology—New York: Columbia University. Subscription $5. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $5. Foreign, $5.25. Editor, Henry T. Moore. Quarterly. 448 pages annually. Abnormal and Social. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (9 Numbers). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

Journal of Applied Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 400 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5. 500 pages annually. Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research.

Genetic Psychology Monographs—Worcester, Mass.: Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Edited by Carl Murchison. Monthly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

(*Continued on Third Cover*)